



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 174198

TO: Sumesh Kaushal
Location: REM-2B85/2C70
Art Unit: 1633
Friday, December 23, 2005
Case Serial Number: 10/787382

From: Toby Port
Location: Biotech-Chem Library
REM-1A59
Phone: 571-272-2523

toby.port@uspto.gov

Search Notes

Examiner Kaushal,

See attached results.

If you have any questions about this search feel free to contact me at any time.

Thank you for using STIC search services!

Toby Port
X22523

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2005, 07:25:38 ; Search time 46 Seconds
(without alignments)
26.959 Million cell updates/sec

Title: US-10-787-382-20

Perfect score: 74

Sequence: 1 FAVENPMRLVAETL 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : Issued Patents AA.*

1: /cgn2_6/ptodata/1/iaa/5_COMB.pep.*

2: /cgn2_6/ptodata/1/iaa/6_COMB.pep.*

3: /cgn2_6/ptodata/1/iaa/H_COMB.pep.*

4: /cgn2_6/ptodata/1/iaa/pCTrus_COMB.pep.*

5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep.*

6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	74	100.0	115	2	US-09-322-409-86
2	74	100.0	115	2	US-09-451-527-86
3	74	100.0	134	2	US-09-322-409-81
4	74	100.0	134	2	US-09-451-527-81
5	74	100.0	134	2	US-09-371-615A-2
6	43	58.1	69	2	US-09-248-796A-22862
7	39	52.7	559	2	US-09-543-681A-7241
8	38	51.4	341	2	US-09-902-540-13873
9	38	51.4	549	2	US-09-134-000C-4375
10	38	51.4	664	2	US-09-591-095-4
11	38	51.4	860	2	US-09-252-991A-26112
12	38	51.4	1205	2	US-09-252-991A-28876
13	37	50.0	133	2	US-09-371-615A-8
14	37	50.0	236	4	PCT-US91-08177-17
15	37	50.0	241	2	US-10-187-790A-16
16	37	50.0	557	2	US-09-949-002-347
17	37	50.0	578	2	US-09-949-002-428
18	37	50.0	586	2	US-08-459-953A-10
19	37	50.0	586	2	US-09-393-212-10
20	36	48.6	95	2	US-08-905-223-307
21	36	48.6	123	2	US-09-583-110-4918
22	36	48.6	196	2	US-09-328-352-4939
23	36	48.6	201	2	US-09-711-164-309
24	36	48.6	201	2	US-09-492-709A-391
25	36	48.6	207	2	US-09-489-039A-10971
26	36	48.6	215	2	US-09-543-681A-7203
27	36	48.6	441	2	US-09-489-039A-13401

28	36	48.6	816	2	US-09-489-039A-10520	Sequence 10520, A
29	36	48.6	2186	1	US-08-822-445-2	Sequence 2, Appli
30	36	48.6	2186	2	US-09-396-540-2	Sequence 2, Appli
31	36	48.6	3672	1	US-08-822-445-12	Sequence 12, Appl
32	36	48.6	3672	2	US-09-396-540-12	Sequence 12, Appl
33	36	48.6	3801	1	US-08-822-445-10	Sequence 10, Appl
34	36	48.6	3801	2	US-09-396-540-10	Sequence 10, Appl
35	35	47.3	75	2	US-09-107-433-3369	Sequence 3369, Ap
36	35	47.3	82	2	US-09-270-767-50829	Sequence 35612, A
37	35	47.3	82	2	US-09-270-767-50829	Sequence 50829, A
38	35	47.3	151	2	US-09-489-039A-11178	Sequence 11178, A
39	35	47.3	151	2	US-09-902-540-14169	Sequence 14169, A
40	35	47.3	159	1	US-08-553-619B-2	Sequence 2, Appli
41	35	47.3	173	2	US-09-248-796A-16707	Sequence 16707, A
42	35	47.3	185	2	US-09-583-110-5154	Sequence 5154, Ap
43	35	47.3	194	2	US-09-107-433-3648	Sequence 3648, Ap
44	35	47.3	197	2	US-09-902-540-14643	Sequence 14643, A
45	35	47.3	201	2	US-09-543-681A-7740	Sequence 7740, Ap
46	35	47.3	205	2	US-10-012-819-196	Sequence 196, App
47	35	47.3	215	2	US-09-543-681A-4324	Sequence 4324, Ap
48	35	47.3	219	1	US-08-855-140-4	Sequence 4, Appli
49	35	47.3	256	1	US-08-211-312-5	Sequence 5, Appli
50	35	47.3	256	2	US-08-472-285-5	Sequence 5, Appli
51	35	47.3	256	2	US-08-472-928-5	Sequence 5, Appli
52	35	47.3	275	2	US-09-107-433-3445	Sequence 3445, Ap
53	35	47.3	282	2	US-09-252-991A-31742	Sequence 31742, A
54	35	47.3	359	2	US-09-602-787A-466	Sequence 466, App
55	35	47.3	359	2	US-09-248-796A-17869	Sequence 17869, A
56	35	47.3	436	2	US-09-248-796A-17505	Sequence 17505, A
57	35	47.3	447	2	US-09-583-110-4550	Sequence 4550, Ap
58	35	47.3	532	2	US-09-902-540-14357	Sequence 14357, A
59	35	47.3	2108	2	US-09-252-991A-31502	Sequence 31502, A
60	34	45.9	358	2	US-09-248-796A-17253	Sequence 17253, A
61	34	45.9	67	1	US-08-530-010-19	Sequence 19, Appl
62	34	45.9	67	1	US-08-484-101B-19	Sequence 19, Appl
63	34	45.9	67	1	US-08-714-524D-19	Sequence 19, Appl
64	34	45.9	110	2	US-09-270-767-62333	Sequence 62333, A
65	34	45.9	112	2	US-10-104-047-3840	Sequence 3840, Ap
66	34	45.9	127	2	US-09-855-323-12	Sequence 12, Appl
67	34	45.9	127	2	US-10-104-047-2959	Sequence 2959, Ap
68	34	45.9	146	2	US-09-270-767-32921	Sequence 32921, A
69	34	45.9	153	2	US-09-270-767-48138	Sequence 48138, A
70	34	45.9	153	1	US-09-902-540-15085	Sequence 15085, A
71	34	45.9	155	1	US-08-324-003A-27	Sequence 27, Appl
72	34	45.9	155	2	US-09-755-836-27	Sequence 27, Appl
73	34	45.9	159	1	US-08-687-559-1	Sequence 1, Appli
74	34	45.9	159	1	US-08-324-003A-22	Sequence 22, Appl
75	34	45.9	159	2	US-09-401-415-1	Sequence 1, Appli
76	34	45.9	159	2	US-09-755-836-22	Sequence 22, Appl
77	34	45.9	159	2	US-09-303-518D-356	Sequence 356, App
78	34	45.9	159	2	US-09-303-518D-358	Sequence 358, App
79	34	45.9	159	2	US-09-303-518D-362	Sequence 362, App
80	34	45.9	159	2	US-09-303-518D-364	Sequence 364, App
81	34	45.9	169	1	US-08-687-559-22	Sequence 22, Appl
82	34	45.9	169	1	US-08-324-003A-17	Sequence 17, Appl
83	34	45.9	169	2	US-09-401-415-22	Sequence 22, Appl
84	34	45.9	169	2	US-09-755-836-17	Sequence 17, Appl
85	34	45.9	174	1	US-08-687-559-21	Sequence 21, Appl
86	34	45.9	174	1	US-08-687-559-23	Sequence 23, Appl
87	34	45.9	174	1	US-08-324-003A-20	Sequence 20, Appl
88	34	45.9	174	2	US-09-401-415-21	Sequence 21, Appl
89	34	45.9	174	2	US-09-401-415-23	Sequence 23, Appl
90	34	45.9	174	2	US-09-755-836-20	Sequence 20, Appl
91	34	45.9	176	2	US-09-134-000C-4221	Sequence 4221, Ap
92	34	45.9	178	1	US-08-324-003A-25	Sequence 25, Appl
93	34	45.9	178	2	US-09-755-836-25	Sequence 25, Appl
94	34	45.9	179	1	US-08-687-559-24	Sequence 24, Appl
95	34	45.9	179	2	US-09-401-415-24	Sequence 24, Appl
96	34	45.9	184	2	US-09-475-316A-99	Sequence 99, Appl
97	34	45.9	184	2	US-09-704-640-99	Sequence 99, Appl
98	34	45.9	213	1	US-08-905-267-13	Sequence 13, Appl
99	34	45.9	213	2	US-09-314-051-13	Sequence 13, Appl
100	34	45.9	213	2	US-09-617-389B-13	Sequence 13, Appl

```
101 34 45.9 214 2 US-09-248-796A-17961 Sequence 17961, A
102 34 45.9 241 4 PCT-US91-08177-7 Sequence 7, Appli
103 34 45.9 249 2 US-09-134-000C-5219 Sequence 5219, Ap
104 34 45.9 254 2 US-09-248-796A-24590 Sequence 24590, A
105 34 45.9 283 2 US-09-314-701-26 Sequence 26, Appli
106 34 45.9 283 2 US-10-314-639-26 Sequence 26, Appli
107 34 45.9 283 2 US-10-059-964A-26 Sequence 26, Appli
108 34 45.9 300 2 US-09-270-767-45716 Sequence 46716, A
109 34 45.9 326 2 US-09-252-991A-18289 Sequence 18289, A
110 34 45.9 345 2 US-09-107-532A-3849 Sequence 3849, Ap
111 34 45.9 349 2 US-09-252-991A-27933 Sequence 27933, A
112 34 45.9 354 1 US-08-843-530B-30 Sequence 30, Appli
113 34 45.9 354 2 US-09-636-728-27 Sequence 27, Appli
114 34 45.9 361 2 US-09-543-681A-7310 Sequence 7310, Ap
115 34 45.9 376 2 US-09-091-725-15 Sequence 15, Appli
116 34 45.9 391 2 US-09-543-681A-5945 Sequence 5945, Ap
117 34 45.9 488 2 US-09-489-039A-13363 Sequence 13363, A
118 34 45.9 494 2 US-09-252-991A-24309 Sequence 24309, A
119 34 45.9 512 2 US-09-107-532A-7248 Sequence 7248, Ap
120 34 45.9 528 2 US-09-543-681A-5551 Sequence 5551, Ap
121 34 45.9 551 2 US-09-252-991A-44209 Sequence 24209, A
122 34 45.9 591 2 US-09-252-991A-27090 Sequence 27090, A
123 34 45.9 641 2 US-09-270-767-44970 Sequence 44970, A
124 34 45.9 714 2 US-09-252-991A-24776 Sequence 24776, A
125 34 45.9 714 2 US-08-969-415-2 Sequence 2, Appli
126 34 45.9 751 2 US-09-538-092-127 Sequence 127, App
127 34 45.9 900 2 US-09-248-796A-18281 Sequence 18281, A
128 34 45.9 947 2 US-09-252-991A-29359 Sequence 29359, A
129 34 45.9 1096 2 US-09-792-024-120 Sequence 120, App
130 34 45.9 1218 2 US-09-252-991A-24869 Sequence 24869, A
131 34 45.9 1890 2 US-09-902-540-15442 Sequence 15442, A
132 34 45.9 3144 1 US-08-246-982A-6 Sequence 6, Appli
133 34 45.9 3144 1 US-08-453-265-6 Sequence 6, Appli
134 34 45.9 3144 1 US-08-457-273B-42 Sequence 42, Appli
135 34 45.9 3144 2 US-08-556-419-21 Sequence 21, Appli
136 34 45.9 3144 2 US-09-041-886-15 Sequence 15, Appli
137 34 45.9 3144 2 US-09-538-092-1118 Sequence 1118, Ap
138 33 5 45.3 138 2 US-09-902-540-14591 Sequence 14591, A
139 33 44.6 36 2 US-09-057-363C-58 Sequence 58, Appli
140 33 44.6 43 2 US-09-265-107-58 Sequence 58, Appli
141 33 44.6 43 2 US-08-379-437-6 Sequence 6, Appli
142 33 44.6 57 1 US-09-492-709A-370 Sequence 370, App
143 33 44.6 103 2 US-08-893-534A-6 Sequence 6, Appli
144 33 44.6 108 2 US-08-996-673-6 Sequence 6, Appli
145 33 44.6 108 2 US-08-939-853A-8 Sequence 8, Appli
146 33 44.6 108 2 US-09-115-395-6 Sequence 6, Appli
147 33 44.6 108 2 US-09-113-977C-35 Sequence 35, Appli
148 33 44.6 108 2 US-09-507-102-6 Sequence 6, Appli
149 33 44.6 108 2 US-09-250-059-6 Sequence 6, Appli
150 33 44.6 108 2 US-09-250-059-6 Sequence 6, Appli
```

ALIGNMENTS

```
RESULT 1
US-09-322-409-86
; Sequence 86, Application US/09322409
; Patent No. 6471957
; BEST LOCAL SIMILARITY 100.0%; Score 74; DB 2; Length 115;
; Mismatches 0; Indels 0; Gaps 0;
; ORGANISM: Canis familiaris
; TYPE: PRT
; LENGTH: 115
; SEQ ID NO 86
;
Query Match 100.0%; Score 74; DB 2; Length 115;
Best Local Similarity 100.0%; Pred. No. 4.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
; ORGANISM: Canis familiaris
; TYPE: PRT
; LENGTH: 115
; SEQ ID NO 86
;
; APPLICANT: Yang, Shumin
; APPLICANT: Dreitz, Matthew J.
; APPLICANT: Wonderling, Ramani S.
; TITLE OF INVENTION: CANINE AND FELINE IMMUNOREGULATORY PROTEINS, NUCLEIC
; FILE REFERENCE: IM-2-C1
; CURRENT APPLICATION NUMBER: US/09/322,409
; CURRENT FILING DATE: 1999-05-28
; EARLIER FILING DATE: 1998-05-29
; NUMBER OF SEQ ID NOS: 154
; SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 81
; LENGTH: 134
; TYPE: PRT
; ORGANISM: Canis familiaris
;
US-09-322-409-81
; Sequence 81, Application US/09322409
; Patent No. 6471957
; BEST LOCAL SIMILARITY 100.0%; Score 74; DB 2; Length 115;
; Mismatches 0; Indels 0; Gaps 0;
; ORGANISM: Canis familiaris
; TYPE: PRT
; LENGTH: 134
; SEQ ID NO 81
;
Query Match 100.0%; Score 74; DB 2; Length 115;
Best Local Similarity 100.0%; Pred. No. 4.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
; ORGANISM: Canis familiaris
; TYPE: PRT
; LENGTH: 134
; SEQ ID NO 81
;
; APPLICANT: Yang, Shumin
; APPLICANT: Dreitz, Matthew J.
; APPLICANT: Wonderling, Ramani S.
; TITLE OF INVENTION: CANINE AND FELINE IMMUNOREGULATORY PROTEINS, NUCLEIC
; FILE REFERENCE: IM-2-C1
; CURRENT APPLICATION NUMBER: US/09/322,409
; CURRENT FILING DATE: 1999-05-28
; EARLIER FILING DATE: 1998-05-29
; NUMBER OF SEQ ID NOS: 154
; SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 81
; LENGTH: 134
; TYPE: PRT
; ORGANISM: Canis familiaris
;
US-09-322-409-81
```


Query Match 100.0%; Score 74; DB 2; Length 134;
Best Local Similarity 100.0%; Pred. No. 5.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FAVENPMNRLVAETL 15
Db 20 FAVENPMNRLVAETL 34

RESULT 4

US-09-451-527-81
; Sequence 81, Application US/09451527
; Patent No. 6482403
; GENERAL INFORMATION:
; APPLICANT: Yang, Gek-kee
; APPLICANT: Yang, Shumin
; APPLICANT: Dreitz, Matthew J.
; APPLICANT: Wonderling, Ramani S.
; TITLE OF INVENTION: CANINE AND FELINE IMMUNOREGULATORY PROTEINS, NUCLEIC
; FILE REFERENCE: IM-2-C2
; CURRENT APPLICATION NUMBER: US/09/451,527
; CURRENT FILING DATE: 1999-12-01
; EARLIER APPLICATION NUMBER: 09/322,409
; EARLIER FILING DATE: 1999-05-28
; EARLIER APPLICATION NUMBER: 60/087,306
; EARLIER FILING DATE: 1998-05-29
; NUMBER OF SEQ ID NOS: 174
; SOFTWARE: Patentin ver. 2.0
; SEQ ID NO 81
; LENGTH: 134
; TYPE: PRT
; ORGANISM: Canis familiaris
US-09-451-527-81

Query Match 100.0%; Score 74; DB 2; Length 134;
Best Local Similarity 100.0%; Pred. No. 5.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FAVENPMNRLVAETL 15
Db 20 FAVENPMNRLVAETL 34

RESULT 5

US-09-371-615A-2
; Sequence 2, Application US/09371615A
; Patent No. 6537781
; GENERAL INFORMATION:
; APPLICANT: IDEXX LABORATORIES
; TITLE OF INVENTION: METHODS AND COMPOSITIONS CONCERNING
; TITLE OF INVENTION: CANINE INTERLEUKIN 5
; FILE REFERENCE: 036040017000US00
; CURRENT APPLICATION NUMBER: US/09/371,615A
; CURRENT FILING DATE: 1999-08-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 134
; TYPE: PRT
; ORGANISM: Canis familiaris
US-09-371-615A-2

Query Match 100.0%; Score 74; DB 2; Length 134;
Best Local Similarity 100.0%; Pred. No. 5.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FAVENPMNRLVAETL 15
Db 20 FAVENPMNRLVAETL 34

RESULT 6

US-09-248-796A-22862
; Sequence 22862, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 22862
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-22862

Query Match 58.1%; Score 43; DB 2; Length 69;
Best Local Similarity 63.6%; Pred. No. 1.2;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 VENPMNRLVAE 13
Db 25 IENPLNRLTAK 35

RESULT 7

US-09-543-681A-7241
; Sequence 7241, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 7241
; LENGTH: 559
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-7241

Query Match 52.7%; Score 39; DB 2; Length 559;
Best Local Similarity 42.9%; Pred. No. 79;
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FAVENPMNRLVAET 14
Db 397 FSLKPNRLINT 410

RESULT 8

US-09-902-540-13873
; Sequence 13873, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiesand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10

; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 13873
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-13873

Query Match 51.4%; Score 38; DB 2; Length 341;
Best Local Similarity 63.6%; Pred. No. 67;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 AVENPMRLVA 12
| | | | |
DB 26 AVQSPDLRLVA 36

RESULT 9
US-09-134-000C-4375
; Sequence 4375, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4375
; LENGTH: 549
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (73)..(73)
; OTHER INFORMATION: Amino acid 73 is Xaa whersin Xaa = any amino acid.
US-09-134-000C-4375

Query Match 51.4%; Score 38; DB 2; Length 549;
Best Local Similarity 57.1%; Pred. No. 1.2e+02;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 AVENPMRLVAETL 15
| | | | |
DB 218 AVDMFINSVEETL 231

RESULT 10
US-09-591-095-4
; Sequence 4, Application US/09591095
; Patent No. 6489461
; GENERAL INFORMATION:
; APPLICANT: Froman, B. & Dehesh, K.
; TITLE OF INVENTION: Nucleic Acid Sequences encoding Proteins
; TITLE OF INVENTION: Involved in Fatty Acid Beta-Oxidation and Methods of Use.
; FILE REFERENCE: 17137/01/US
; CURRENT APPLICATION NUMBER: US/09/591,095
; CURRENT FILING DATE: 2000-06-08
; EARLIER APPLICATION NUMBER: US 60/138,162
; EARLIER FILING DATE: 1999-06-08
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 664
; TYPE: PRT
; ORGANISM: Arabidopsis sp.
US-09-591-095-4

Query Match 51.4%; Score 38; DB 2; Length 664;
Best Local Similarity 60.0%; Pred. No. 1.5e+02;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 PAVENPMRLVAETL 15
| | | | |
DB 32 FDSVNSRLVANDL 46

RESULT 11
US-09-252-991A-26112
; Sequence 26112, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26112
; LENGTH: 860
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26112

Query Match 51.4%; Score 38; DB 2; Length 860;
Best Local Similarity 58.3%; Pred. No. 2e+02;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 4 ENPMNRLVAETL 15
| | | | |
DB 744 DRPNQLVAKGL 755

RESULT 12
US-09-252-991A-28876
; Sequence 28876, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28876
; LENGTH: 1205
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28876

Query Match 51.4%; Score 38; DB 2; Length 1205;
Best Local Similarity 61.5%; Pred. No. 3e+02;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 VENPMNRLVAETL 15
| | | | |
DB 964 VDNRLRLQAEYL 976

RESULT 13
US-09-371-615A-8

DD 296 NFM

RESULT 17
US-09-949-002-428
; Sequence 428, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 428
; LENGTH: 578
; TYPE: PRT
; ORGANISM: Human
US-09-949-002-428

Query Match 50.0%; Score 37; DB 2; Length 578;
Best Local Similarity 77.8%; Pred. No. 1.9e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 NPMNRLVAE 13
|||:||||
DB 325 NPMNRLTAE 333

RESULT 18
US-08-459-953A-10
; Sequence 10, Application US/08459953A
; Patent No. 6030822
; GENERAL INFORMATION:
; APPLICANT: Lechner, Cornelia
; APPLICANT: Moller, Niels P.H.
; APPLICANT: Ullrich, Axel
; TITLE OF INVENTION: EXTRACELLULAR SIGNAL-RELATED
; TITLE OF INVENTION: KINASE, SEQUENCES, AND
; TITLE OF INVENTION: METHODS OF PRODUCTION
; TITLE OF INVENTION: AND USE
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; STATE: Los Angeles
; COUNTRY: California
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,953A
; FILING DATE: June 2, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/029,494
; FILING DATE: March 19, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 209/267
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440

TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 10 :
; SEQUENCE CHARACTERISTICS:
; LENGTH: 586 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-459-953A-10

Query Match 50.0%; Score 37; DB 2; Length 586;
Best Local Similarity 77.8%; Pred. No. 1.9e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 NPMNRLVAE 13
|||:||||
DB 325 NPMNRLTAE 333

RESULT 19
US-09-393-212-10
; Sequence 10, Application US/09393212
; Patent No. 6579972
; GENERAL INFORMATION:
; APPLICANT: Lechner, Cornelia
; APPLICANT: Moller, Niels P.H.
; APPLICANT: Ullrich, Axel
; TITLE OF INVENTION: EXTRACELLULAR SIGNAL-RELATED
; TITLE OF INVENTION: KINASE, SEQUENCES, AND
; TITLE OF INVENTION: METHODS OF PRODUCTION
; AND USE
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; STATE: Los Angeles
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/393,212
; FILING DATE: 09-Sep-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/459,953A
; FILING DATE: June 2, 1995
; APPLICATION NUMBER: 08/029,494
; FILING DATE: March 19, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 209/267
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 10 :
; SEQUENCE CHARACTERISTICS:
; LENGTH: 586 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-393-212-10

Query Match 50.0%; Score 37; DB 2; Length 586;
Best Local Similarity 77.8%; Pred. No. 1.9e+02;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 5 NPMNRLVAE 13
|||:||||
Db 325 NPMNRLTAE 333

RESULT 20

US-08-905-223-307
; Sequence 307, Application US/08905223
; Patent No. 6222029

GENERAL INFORMATION:

; APPLICANT: Edwards, Jean-Baptiste D.
; APPLICANT: Duclert, Aymeric

; APPLICANT: Lacroix, Bruno

; TITLE OF INVENTION: 5' ESTs FOR SECRETED PROTEINS

; NUMBER OF SEQUENCES: 503

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Knobbe, Martens, Olson & Bear

; STREET: 501 West Broadway

; CITY: San Diego

; STATE: California

; COUNTRY: USA

; ZIP: 92101-3505

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy Disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: Win95

; SOFTWARE: Word

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/905,223

; FILING DATE:

; CLASSIFICATION: 536

; ATTORNEY/AGENT INFORMATION:

; NAME: Israelsen, Ned A.

; REGISTRATION NUMBER: 29,655

; REFERENCE/DOCKET NUMBER:

; TELEPHONE: (619) 235-8550

; TELEFAX: (619) 235-0176

; INFORMATION FOR SEQ ID NO: 307:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 95 amino acids

; TYPE: AMINO ACID

; TOPOLOGY: LINEAR

; MOLECULE TYPE: PROTEIN

; ORIGINAL SOURCE:

; ORGANISM: Homo Sapiens

; TISSUE TYPE: Brain

; FEATURE:

; NAME/KEY: sig_peptide

; LOCATION: -92--1

; IDENTIFICATION METHOD: Von Heijne matrix

; OTHER INFORMATION: score 6.8

; OTHER INFORMATION: seq ILIASSLPTLSHP/AP

US-08-905-223-307

Query Match 48.6%; Score 36; DB 2; Length 95;

Best Local Similarity 80.0%; Pred. No. 34;

Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 5 NPMNRLVAET 14
|||||
Db 53 NPMNRLVAGT 62

RESULT 21

US-09-583-110-4918

; Sequence 4918, Application US/09583110

; Patent No. 6699703

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al.

; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus

; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics

; FILE REFERENCE: PATH00-07A

; CURRENT APPLICATION NUMBER: US/09/583,110

; CURRENT FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: US 09/107,433

; PRIOR FILING DATE: 1998-06-30

; PRIOR APPLICATION NUMBER: US 60/085,131

; PRIOR FILING DATE: 1998-05-12

; PRIOR APPLICATION NUMBER: US 60/051,553

; PRIOR FILING DATE: 1997-07-02

; NUMBER OF SEQ ID NOS: 5322

; SEQ ID NO 4918

; LENGTH: 123

; TYPE: PRT

; ORGANISM: Streptococcus pneumoniae

US-09-583-110-4918

Query Match 48.6%; Score 36; DB 2; Length 123;

Best Local Similarity 53.3%; Pred. No. 47;

Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 FAVENPMNRLVAETL 15
|||:||||
Db 41 FGVEASMGRIYPETL 55

RESULT 22

US-09-328-352-4939

; Sequence 4939, Application US/09328352

; Patent No. 6562958

; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

; FILE REFERENCE: GTC99-03PA

; CURRENT APPLICATION NUMBER: US/09/328,352

; CURRENT FILING DATE: 1999-06-04

; NUMBER OF SEQ ID NOS: 8252

; SEQ ID NO 4939

; LENGTH: 196

; TYPE: PRT

; ORGANISM: Acinetobacter baumannii

US-09-328-352-4939

Query Match 48.6%; Score 36; DB 2; Length 196;

Best Local Similarity 42.9%; Pred. No. 81;

Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 FAVENPMNRLVAET 14
|||:||||
Db 21 YALEMPMNAVATDT 34

RESULT 23

US-09-711-164-309

; Sequence 309, Application US/09711164

; Patent No. 6589738

; GENERAL INFORMATION:

; APPLICANT: Forsyth, R. Allyn

; APPLICANT: Ohlsen, Kari

; APPLICANT: Zyskind, Judith

; TITLE OF INVENTION: GENES ESSENTIAL FOR MICROBIAL PROLIFERATION AND ANTISENSE THEREOF

; FILE REFERENCE: ELITRA.008A

; CURRENT APPLICATION NUMBER: US/09/711,164

; CURRENT FILING DATE: 2000-11-09

; PRIOR APPLICATION NUMBER: US 60/164415

; PRIOR FILING DATE: 1999-11-9

; NUMBER OF SEQ ID NOS: 469

; SOFTWARE: Fast-Seq for Windows Version 4.0

; SEQ ID NO 309

; LENGTH: 201

; TYPE: PRT

; ORGANISM: Escherichia coli

US-09-711-164-309

Query Match 48.6%; Score 36; DB 2; Length 201;
Best Local Similarity 46.7%; Pred. No. 84;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 FAVENPMNRLVAETL 15
|:|:| |:|:| |:|:|
Db 124 FSVEAPKTKLLAQKL 138

RESULT 24

US-09-492-709A-391
; Sequence 391, Application US/09492709A
; Patent No. 6720139
; GENERAL INFORMATION:
; APPLICANT: Zyskind, Judith
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Trawick, John
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Froelich, Jamie M.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION IN
; TITLE OF INVENTION: ESCHERICHIA COLI
; FILE REFERENCE: ELITRA.001A
; CURRENT APPLICATION NUMBER: US/09/492,709A
; CURRENT FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 485
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 391
; LENGTH: 201
; TYPE: PRT
; ORGANISM: E. Coli
US-09-492-709A-391

Query Match 48.6%; Score 36; DB 2; Length 201;
Best Local Similarity 46.7%; Pred. No. 84;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 FAVENPMNRLVAETL 15
|:|:| |:|:| |:|:|
Db 124 FSVEAPKTKLLAQKL 138

RESULT 25

US-09-489-039A-10971
; Sequence 10971, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10971
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10971

Query Match 48.6%; Score 36; DB 2; Length 207;
Best Local Similarity 46.7%; Pred. No. 87;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 FAVENPMNRLVAETL 15
|:|:| |:|:| |:|:|
Db 130 FSVEAPKTKLLAQKL 144

RESULT 26

US-09-543-681A-7203
; Sequence 7203, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 7203
; LENGTH: 215
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-7203

Query Match 48.6%; Score 36; DB 2; Length 215;
Best Local Similarity 46.7%; Pred. No. 91;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 FAVENPMNRLVAETL 15
|:|:| |:|:| |:|:|
Db 138 FSVEAPKTKLLAQKL 152

RESULT 27

US-09-489-039A-13401
; Sequence 13401, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13401
; LENGTH: 441
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13401

Query Match 48.6%; Score 36; DB 2; Length 441;
Best Local Similarity 58.3%; Pred. No. 2.1e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 VENPMNRLVAET 14
|:|:| |:|:| |:|:|
Db 160 VENPOHRAAAT 171

RESULT 28

US-09-489-039A-10520
; Sequence 10520, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29

; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10520
; LENGTH: 816
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10520

Query Match 48.6%; Score 36; DB 2; Length 816;
Best Local Similarity 57.1%; Pred. No. 4.4e+02;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 2 AVENPMRLVAETL 15
||| ||| ||| |||
Db 360 AVEQMNELLSEYL 373

RESULT 29

US-08-822-445-2
; Sequence 2, Application US/08822445
; Patent No. 5952223

; GENERAL INFORMATION:

; APPLICANT: Kaplan, Jerry
; APPLICANT: Perou, Charles
; APPLICANT: Moore, Karen
; TITLE OF INVENTION: COMPOSITIONS FOR THE DIAGNOSIS
; AND TREATMENT OF CHEDIAK-HIGASHI SYNDROME
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036/2711

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/822,445
; FILING DATE: 21-MAR-1997

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-062-999

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIE

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 2186 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal

US-08-822-445-2

Query Match 48.6%; Score 36; DB 1; Length 2186;
Best Local Similarity 50.0%; Pred. No. 1.4e+03;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 2 AVENPMRLVAETL 15
||| ||| ||| |||
Db 1775 AVEQVQRRALETM 1788

RESULT 30

US-09-396-540-2

; Sequence 2, Application US/09396540
; Patent No. 6310182

; GENERAL INFORMATION:
; APPLICANT: Kaplan, Jerry
; APPLICANT: Perou, Charles
; APPLICANT: Moore, Karen
; TITLE OF INVENTION: COMPOSITIONS FOR THE DIAGNOSIS
; AND TREATMENT OF CHEDIAK-HIGASHI SYNDROME
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036/2711

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/396,540
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/822,445
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-062-999

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIE

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 2186 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal

US-09-396-540-2

Query Match 48.6%; Score 36; DB 2; Length 2186;
Best Local Similarity 50.0%; Pred. No. 1.4e+03;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 2 AVENPMRLVAETL 15
||| ||| ||| |||
Db 1775 AVEQVQRRALETM 1788

Search completed: December 17, 2005, 07:34:27
Job time : 48 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2005, 07:07:52 ; Search time 186 Seconds
(without alignments)
35.434 Million cell updates/sec

Title: US-10-787-382-20

Perfect score: 74

Sequence: 1 FAVENPMRLVAETL 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

A_Geneseq_21.*

1: Geneseqp1980s.*

2: Geneseqp1990s.*

3: Geneseqp2000s.*

4: Geneseqp2001s.*

5: Geneseqp2002s.*

6: Geneseqp2003as.*

7: Geneseqp2003bs.*

8: Geneseqp2004s.*

9: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	74	100.0	84	4	AAB72616 Canine in
2	74	100.0	115	3	AAY58220 Canine ma
3	74	100.0	134	3	AAY58219 Canine in
4	74	100.0	134	4	AAB72615 Canine in
5	55	74.3	132	2	AAW08479 Ovine IL-
6	46	62.2	427	4	ABG99883 S. cinna
7	43	58.1	159	3	ABG99883 Pepper mi
8	41	55.4	307	4	ABB71894 Drosophil
9	41	55.4	307	8	AD08016 Fly poly
10	41	55.4	514	8	ADS27735 Bacterial
11	40	54.1	59	8	ADU47425 Yarrowia
12	39	52.7	518	8	ABG69884 Drosophil
13	39	52.7	559	7	ADP06956 Bacterial
14	39	52.7	756	8	ADJ49966 Oil-assoc
15	39	52.7	804	4	AU34800 E. coli c
16	39	52.7	804	6	ABU28825 Protein e
17	38.5	52.0	711	8	ADN25977 Bacterial
18	38	51.4	59	5	ABP06885 Human ORF
19	38	51.4	252	6	ABU18938 Protein e
20	38	51.4	341	9	ABM94674 M. xanthu
21	38	51.4	549	7	ADH86490 Enterococ
22	38	51.4	664	4	AAE20000 Arabidops
23	38	51.4	860	7	ABO77366 Pseudomon
24	38	51.4	956	4	ABB53002 Escherich

25	38	51.4	1205	7	ABO80130	Ab080130 Pseudomon
26	37.5	50.7	324	9	AEA62655	Aea62655 Mitochond
27	37	50.0	88	5	ABP02493	ABP02493 Human ORF
28	37	50.0	115	2	AAW05273	Aaw05273 N-termina
29	37	50.0	115	2	AAW72948	Aaw72948 Truncated
30	37	50.0	133	1	AAW71064	Aaw71064 Murine eo
31	37	50.0	133	1	AAW82969	Aaw82969 B cell di
32	37	50.0	133	1	AAW80280	AAW80280 Murine bl
33	37	50.0	133	2	AAW96963	AAW96963 T cell re
34	37	50.0	133	2	AAW72949	AAW72949 T cell re
35	37	50.0	133	2	AAW72947	AAW72947 T cell re
36	37	50.0	133	4	AAW72618	AAW72618 Murine in
37	37	50.0	134	5	ABG94349	ABG94349 Mouse C-I
38	37	50.0	134	5	ABG80661	ABG80661 Mouse IL-
39	37	50.0	134	7	ADK17220	ADK17220 Mouse C-I
40	37	50.0	162	8	ADN47619	ADN47619 Thermococ
41	37	50.0	236	2	AAW24191	AAW24191 Bovine RS
42	37	50.0	246	9	ADW17740	ADW17740 Pinus rad
43	37	50.0	311	6	ABU28549	ABU28549 Protein e
44	37	50.0	311	6	ABU45268	ABU45268 Protein e
45	37	50.0	360	7	ADC61397	ADC61397 Nicotiana
46	37	50.0	448	4	AAW25874	AAW25874 Human pro
47	37	50.0	557	7	ADF45053	ADF45053 Human kin
48	37	50.0	557	8	ADS92972	ADS92972 Mitogen-a
49	37	50.0	558	8	ADS10754	ADS10754 Human the
50	37	50.0	583	9	ADV97853	ADV97853 Murine pr
51	37	50.0	586	7	ADC64460	ADC64460 Human ext
52	37	50.0	586	8	ADG98196	ADG98196 Human ext
53	37	50.0	586	8	ABM84222	ABM84222 Human dia
54	37	50.0	587	6	AAE32021	AAE32021 Human kin
55	37	50.0	587	6	ADS12008	ADS12008 Human the
56	37	50.0	603	4	ABG29655	ABG29655 Novel hum
57	37	50.0	606	4	ABB69121	ABB69121 Drosophil
58	37	50.0	615	8	ADS27325	ADS27325 Bacterial
59	37	50.0	663	8	ADS26595	ADS26595 Bacterial
60	37	50.0	663	8	ADS26970	ADS26970 Bacterial
61	37	50.0	985	4	AAW90623	AAW90623 C Glutam
62	37	50.0	1021	3	AAW77728	AAW77728 Human G p
63	37	50.0	1021	3	ADE55164	ADE55164 Human Pro
64	37	50.0	1021	7	ADE55168	ADE55168 Human Pro
65	37	50.0	1021	7	ADE55172	ADE55172 Human Pro
66	37	50.0	1021	7	ADE55176	ADE55176 Human Pro
67	37	50.0	1027	7	ABM85538	ABM85538 Mouse pro
68	37	50.0	1155	4	AAW92010	AAW92010 C Glutam
69	37	50.0	1155	7	ADJ87397	ADJ87397 DNA repli
70	37	50.0	1155	9	ABE13129	ABE13129 C. glutam
71	37	50.0	1351	4	ABE61727	ABE61727 Drosophil
72	37	50.0	1459	7	ADC26275	ADC26275 Human NOV
73	37	50.0	1478	5	AAU10540	AAU10540 Rat CIRL-
74	37	50.0	1478	7	ADD46680	ADD46680 Rat Prote
75	37	50.0	1488	7	ADE55162	ADE55162 Rat Prote
76	37	50.0	1488	7	ADE55174	ADE55174 Rat Prote
77	37	50.0	1488	7	ADE55170	ADE55170 Rat Prote
78	37	50.0	1488	7	ADE55166	ADE55166 Rat Prote
79	37	50.0	1488	8	ADM97582	ADM97582 Human cal
80	37	50.0	1878	4	ABB61852	ABB61852 Drosophil
81	37	50.0	2092	6	ABJ25403	ABJ25403 Aspergill
82	37	50.0	2092	6	ABJ26003	ABJ26003 Aspergill
83	36.5	49.3	288	8	ADT59574	ADT59574 Plant pol
84	36	48.6	77	4	ABB41650	ABB41650 Peptide #
85	36	48.6	77	4	AAW35444	AAW35444 Peptide #
86	36	48.6	77	4	AAW75331	AAW75331 Human bon
87	36	48.6	77	4	AAW62520	AAW62520 Human bra
88	36	48.6	77	4	ABG57088	ABG57088 Human liv
89	36	48.6	95	2	AAW13023	AAW13023 Human sec
90	36	48.6	123	8	ADK48403	ADK48403 Streptoco
91	36	48.6	146	4	AAU03698	AAU03698 Group B S
92	36	48.6	155	4	ABG07176	ABG07176 Novel hum
93	36	48.6	175	3	AAW53339	AAW53339 Human col
94	36	48.6	196	6	ADA33652	ADA33652 Acinetoba
95	36	48.6	197	4	AAW73391	AAW73391 Human col
96	36	48.6	200	6	ABU45632	ABU45632 Protein e
97	36	48.6	201	3	AAW16033	AAW16033 E. coli p

98	36	48.6	201	4	AAU34765
99	36	48.6	201	4	AAU38380
100	36	48.6	201	4	AAU38380
101	36	48.6	201	4	AAU38380
102	36	48.6	201	6	AAU38380
103	36	48.6	201	6	AAU38380
104	36	48.6	201	6	AAU38380
105	36	48.6	201	6	AAU38380
106	36	48.6	201	6	AAU38380
107	36	48.6	201	6	AAU38380
108	36	48.6	201	7	AAU38380
109	36	48.6	215	7	AAU38380
110	36	48.6	220	8	AAU38380
111	36	48.6	221	8	AAU38380
112	36	48.6	254	4	AAU38380
113	36	48.6	255	4	AAU38380
114	36	48.6	255	5	AAU38380
115	36	48.6	257	6	AAU38380
116	36	48.6	258	3	AAU38380
117	36	48.6	266	8	AAU38380
118	36	48.6	309	4	AAU38380
119	36	48.6	346	8	AAU38380
120	36	48.6	363	4	AAU38380
121	36	48.6	375	6	AAU38380
122	36	48.6	418	6	AAU38380
123	36	48.6	441	7	AAU38380
124	36	48.6	477	6	AAU38380
125	36	48.6	486	5	AAU38380
126	36	48.6	513	7	AAU38380
127	36	48.6	571	5	AAU38380
128	36	48.6	571	8	AAU38380
129	36	48.6	571	8	AAU38380
130	36	48.6	571	8	AAU38380
131	36	48.6	603	6	AAU38380
132	36	48.6	641	6	AAU38380
133	36	48.6	654	5	AAU38380
134	36	48.6	717	6	AAU38380
135	36	48.6	730	9	AAU38380
136	36	48.6	803	4	AAU38380
137	36	48.6	804	4	AAU38380
138	36	48.6	804	6	AAU38380
139	36	48.6	816	7	AAU38380
140	36	48.6	1079	4	AAU38380
141	36	48.6	1175	8	AAU38380
142	36	48.6	1447	8	AAU38380
143	36	48.6	2186	2	AAU38380
144	36	48.6	3672	2	AAU38380
145	36	48.6	3788	2	AAU38380
146	36	48.6	3801	2	AAU38380
147	36	48.6	3801	8	AAU38380
148	35	47.3	37	7	AAU38380
149	35	47.3	59	8	AAU38380
150	35	47.3	75	8	AAU38380

ALIGNMENTS

RESULT 1
 AAB72616
 ID AAB72616 standard; protein; 84 AA.
 XX
 AC AAB72616;
 XX
 DT 04-MAY-2001 (first entry)
 XX
 DE Canine interleukin-5 protein #2.
 XX
 KW Dog; interleukin-5; IL-5; allergy; cancer; gene therapy;
 XX inflammatory reaction.
 XX
 OS Canis sp.
 XX

AAU34765 E. coli c
 AAU38380 Salmonell
 AAG98839 E. coli g
 AAG98839 E. coli g
 ABU50212 Protein e
 ABU40982 Protein e
 ABU31401 Protein e
 ABU47784 Protein e
 ABU27723 Protein e
 ABU4733 Protein e
 ABU64454 Klebsiell
 ADF06918 Bacteri
 ADV06460 Bacteri
 ADM74015 Thale cre
 ABU06072 N. mening
 AAE04891 Human tra
 ABP51810 Human pot
 ABP78820 N. gonorr
 AAB41622 Human ORF
 ADN19554 Bacteri
 ABG17090 Novel hum
 ADY08364 Plant ful
 ADG27800 Human nov
 ADB07634 Allolococ
 ABU31504 Protein e
 ABO66884 Klebsiell
 ABU23254 Protein e
 ABB47843 Listeria
 ADO55911 Bacillus
 ABP25425 Streptoco
 ADV8216 Streptoco
 ADV79469 Streptoco
 ADV81641 Streptoco
 ABU45042 Protein e
 ABJ25373 Aspergill
 ABB92841 Herbicida
 ABJ25973 Aspergill
 AEA49153 L. rhamo
 ABU28371 Protein e
 AAU38327 Salmonell
 ABU48260 Protein e
 ABO64003 Klebsiell
 ABB71337 Drosophil
 ADS44237 Bacteri
 ADN2331 Bacteri
 AAW31948 Mouse 22B
 AAW31950 Human bg
 AAW23594 Murine Ly
 AAW31949 Human bg
 ADRI4706 Human NF-
 ADD35447 Streptoco
 ADU47422 Saccharom
 ADR94734 Novel S.

PN WO200111049-A2.
 XX 15-FEB-2001.
 XX 09-AUG-2000; 2000WO-US021651.
 XX 10-AUG-1999; 99US-00371615.
 XX (IDEX-) IDEXX LAB INC.
 XX Guo H, Lawton R, Mermer B, Aiyappa AP;
 PI WPI; 2001-191542/19.
 XX N-PSDB; AAF74305.
 XX Novel canine interleukin 5 polynucleotide and polypeptides are used for
 XX generating antibodies which are useful in treating allergies in dogs.
 XX Example 1; Fig 1; 48pp; English.
 XX The present invention provides the protein and coding sequences of the
 XX canine interleukin-5 (IL-5) protein. This can be used to treat allergies,
 XX cancer and inflammatory reactions in dogs. The present sequence is one
 XX version of the IL-5 protein shown in the specification
 XX Sequence 84 AA;
 SQ

Query Match 100.0%; Score 74; DB 4; Length 84;
 Best Local Similarity 100.0%; Pred. No. 5.5e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FAVENPMRLVAETL 15
 DB 20 FAVENPMRLVAETL 34
 |||||
 |||||

RESULT 2
 AAY58220
 ID AAY58220 standard; protein; 115 AA.
 XX
 AC AAY58220;
 XX 14-MAR-2000 (first entry)
 XX Canine mature interleukin-5 (IL-5).
 DE Interleukin-5; IL-5; antibody; canine; inhibitor; immune response;
 XX immunoregulation; tumour; cancer; autoimmune disease; vaccine.
 KW Canis familiaris.
 OS
 XX WO9961618-A2.
 XX 02-DEC-1999.
 XX 28-MAY-1999; 99WO-US011942.
 XX 29-MAY-1998; 98US-0087306P.
 XX (HESK-) HESKA CORP.
 XX Sim G, Yang S, Dreitz MJ, Wonderling RS;
 PI WPI; 2000-072623/06.
 XX N-PSDB; AAZ55550, AAZ55551.
 XX Nucleic acids encoding immunoregulatory proteins from cats or dogs,
 XX useful for treating or preventing e.g. tumors or autoimmune disease.
 XX Claim 3h; Page 227; 264pp; English.
 XX Sequences AAY58219-Y58220 represent encoded and mature canine interleukin
 XX -5 (IL-5). The invention relates to canine IL-4, canine or feline Fit-3
 CC

CC ligand, canine or feline CD40, canine or feline CD154 (CD40 ligand),
 CC canine IL-5, canine IL-13, feline interferon-alpha (IFN-alpha) and feline
 CC granulocyte macrophage colony-stimulating factor (GM-CSF), and nucleotides
 CC which encode these immunoregulatory proteins. The proteins, their
 CC associated nucleic acids, specific antibodies and inhibitors may be used
 CC as vaccines for therapeutic or prophylactic regulation of an immune
 CC response in animals (particularly cats, dogs, horses and humans). They
 CC may be used to treat autoimmune or infectious diseases including
 CC allergies, tumours, inflammation and graft rejection, and to increase the
 CC response from a co-administered antigen. The nucleotide sequences can
 CC also be used for the recombinant production of a protein, while
 CC nucleotide fragments are useful as probes, as amplification primers and
 CC as sources of inhibitory therapeutics (e.g., antisense oligonucleotides).
 CC The proteins may be used to raise antibodies and to screen for modulators
 CC of activity, while the antibodies may be used in detection, and in drug
 CC targeting
 CC
 CC Sequence 115 AA;
 SQ
 Query Match 100.0%; Score 74; DB 3; Length 115;
 Best Local Similarity 100.0%; Pred. No. 8.2e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 FAVENPMNRLVAETL 15
 Db 1 FAVENPMNRLVAETL 15
 RESULT 3
 AAY58219
 ID AAY58219 standard; protein; 134 AA.
 XX
 AC AAY58219;
 DT 14-MAR-2000 (first entry)
 XX
 DE Canine interleukin-5 (IL-5).
 XX
 KW Interleukin-5; IL-5; antibody; canine; inhibitor; immune response;
 KW immunoregulation; tumour; cancer; autoimmune disease; vaccine.
 XX
 OS Canis familiaris.
 XX
 FN WO9961618-A2.
 XX
 PD 02-DEC-1999.
 XX
 PF 28-MAY-1999; 98WO-US011942.
 XX
 PR 29-MAY-1998; 98US-0087306P.
 XX
 PA (HESK-) HESKA CORP.
 XX
 PI Sim G, Yang S, Dreitz MJ, Wonderling RS;
 XX
 DR WPI; 2000-072623/06.
 DR N-PSDB; AAZ55546, AAZ55547, AAZ55548, AAZ55549.
 XX
 PT Nucleic acids encoding immunoregulatory proteins from cats or dogs,
 PT useful for treating or preventing e.g. tumors or autoimmune disease.
 XX
 XX Claim 3h; Page 224; 264pp; English.
 PS
 XX Sequences AAY58219-Y58220 represent encoded and mature canine interleukin
 CC -5 (IL-5). The invention relates to canine IL-4, canine or feline Flt-3
 CC ligand, canine or feline CD40, canine or feline CD154 (CD40 ligand),
 CC canine IL-5, canine IL-13, feline interferon-alpha (IFN-alpha) and feline
 CC granulocyte macrophage colony-stimulating factor (GM-CSF), and nucleotides
 CC which encode these immunoregulatory proteins. The proteins, their
 CC associated nucleic acids, specific antibodies and inhibitors may be used
 CC as vaccines for therapeutic or prophylactic regulation of an immune
 CC response in animals (particularly cats, dogs, horses and humans). They
 CC may be used to treat autoimmune or infectious diseases including

CC allergies, tumours, inflammation and graft rejection, and to increase the
 CC response from a co-administered antigen. The nucleotide sequences can
 CC also be used for the recombinant production of a protein, while
 CC nucleotide fragments are useful as probes, as amplification primers and
 CC as sources of inhibitory therapeutics (e.g., antisense oligonucleotides).
 CC The proteins may be used to raise antibodies and to screen for modulators
 CC of activity, while the antibodies may be used in detection, and in drug
 CC targeting
 CC
 CC Sequence 134 AA;
 SQ
 Query Match 100.0%; Score 74; DB 3; Length 134;
 Best Local Similarity 100.0%; Pred. No. 9.9e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 FAVENPMNRLVAETL 15
 Db 20 FAVENPMNRLVAETL 34
 RESULT 4
 AAB72615
 ID AAB72615 standard; protein; 134 AA.
 XX
 AC AAB72615;
 DT 04-MAY-2001 (first entry)
 XX
 DE Canine interleukin-5 protein #1.
 XX
 KW Dog; interleukin-5; IL-5; allergy; cancer; gene therapy;
 KW inflammatory reaction.
 XX
 OS Canis sp.
 XX
 FN WO200111049-A2.
 XX
 PD 15-FEB-2001.
 XX
 PF 09-AUG-2000; 2000WO-US021651.
 XX
 PR 10-AUG-1999; 99US-00371615.
 XX
 PA (IDEX-) IDEXX LAB INC.
 XX
 PI Guo H, Lawton R, Mermer B, Aiyappa AP;
 XX
 DR WPI; 2001-191542/19.
 DR N-PSDB; AAF74300.
 XX
 PT Novel canine interleukin 5 polynucleotide and polypeptides are used for
 PT generating antibodies which are useful in treating allergies in dogs.
 XX
 PS Claim 29; Page 46-47; 48pp; English.
 XX
 XX The present invention provides the protein and coding sequences of the
 CC canine interleukin-5 (IL-5) protein. This can be used to treat allergies,
 CC cancer and inflammatory reactions in dogs. The present sequence is one
 CC version of the IL-5 protein shown in the specification
 CC
 CC Sequence 134 AA;
 SQ
 Query Match 100.0%; Score 74; DB 4; Length 134;
 Best Local Similarity 100.0%; Pred. No. 9.9e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 FAVENPMNRLVAETL 15
 Db 20 FAVENPMNRLVAETL 34
 RESULT 5
 AAW08479

ID AAW08479 standard; protein; 132 AA.
 XX AC AAW08479;
 AC
 DT 17-OCT-2003 (revised)
 DT 24-SEP-1997 (first entry)
 DT
 XX Ovine IL-5.
 DE
 XX Cytokine; ovine; sheep; interleukin-5; interleukin-12; IL-5; IL-12;
 KW livestock; cow; stress; transport; vaccine adjuvant; veterinary; cancer;
 KW immunosuppression; allergy; reproductive system; growth; early maturity;
 KW antibody; diagnosis; immunopotentiator;
 KW early haematopoietic progenitor cell; cytotoxic cell; thymocyte;
 KW secretion; Igm; IgA; bacterial endotoxin; gamma-interferon.
 XX
 OS Ovis aries.
 XX
 XX WO9700321-A1.
 PN
 XX 03-JAN-1997.
 PD
 XX
 XX 14-JUN-1996; 96WO-AU000360.
 PF
 XX 14-JUN-1995; 95AU-00003502.
 PR
 XX 27-OCT-1995; 95AU-00006244.
 PR
 XX (CSIR) COMMONWEALTH SCI & IND RES ORG.
 PA
 XX
 XX Seow H, Wood P;
 PI
 XX WPI, 1997-077528/07.
 DR
 DR N-PSDB; AAT50755; AAT50756.
 DR
 XX Nucleic acid encoding ovine interleukin-5 or -12 - used as vaccine
 PT adjuvants and to treat or prevent microbial infections in livestock.
 PT
 XX Claim 31; Page 39-40; 78pp; English.
 PS
 XX This protein sequence represents ovine interleukin-5 (IL-5). Ovine IL-5
 CC or IL-12 are used to treat and/or prevent infections in livestock (esp.
 CC cows and sheep), particularly where the animals are stressed, e.g. during
 CC transport. IL-5 and IL-12 can also be used as adjuvants in vaccines for
 CC veterinary use (partic. weakly immunogenic subunit or synthetic peptide
 CC vaccines). They may also be used to treat cancer, immunosuppression and
 CC allergy, to enhance/suppress the reproductive system and to promote
 CC growth or early maturity. Optionally interleukin can be delivered from
 CC constructs or delivery cells and antibodies are useful in enzyme
 CC immunoassays for rapid diagnosis of infection. The interleukins are
 CC immunopotentiators, especially IL-5 promotes growth of early
 CC haematopoietic progenitor cells and generation of cytotoxic cells from
 CC thymocytes, also it stimulates production and secretion of Igm and IgA
 CC (in synergism with bacterial endotoxin). IL-12 induces production of
 CC gamma-interferon by, and proliferation of, T and NK cells and increases
 CC the (non-)specific cytolytic lymphocyte response. The genetic constructs
 CC can also be used for in vitro production of IL-5 or -12. (Updated on 17-
 CC OCT-2003 to standardise OS field)
 XX
 XX Sequence 132 AA;
 SQ
 Query Match 74.3%; Score 55; DB 2; Length 132;
 Best Local Similarity 85.7%; Pred. No. 0.04;
 Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 2 AVENPMNRLVAETL 15
 DB 19 AVESTMNLVAETL 32
 |||: |||||
 RESULT 6
 ABG99883
 ID ABG99883 standard; protein; 427 AA.
 XX
 AC ABG99883;
 DT 16-JAN-2003 (first entry)
 DT
 XX S. cinamonensis CapK homologue.
 DE
 XX Monensin; gene cluster; polyketide synthase; antibiotic; antihelminthic;
 KW insecticide; immunosuppressant; antifungal; antibacterial; polyether;
 KW mon BI; mon BII; mon CI; mon CII; mon H; mon RI; mon RII; mon T; mon AIX;
 KW mon AX.
 XX Streptomyces cinamonensis.
 OS
 XX WO200168867-A1.
 PN
 XX 20-SEP-2001.
 PD
 XX 30-MAY-2000; 2000WO-GB002072.
 PF
 XX 28-MAY-1999; 99GB-00012563.
 PR
 XX (BIOT-) BIOTICA TECHNOLOGY LTD.
 PA
 XX Leadlay PF, Staunton J, Oliynyk M;
 PI
 XX WPI, 2001-611393/70.
 DR
 DR N-PSDB; ABX04971.
 DR
 XX New DNA sequence encoding polyketide synthase, useful for the production
 PT of polyketides such as antibiotic monensin.
 PT
 XX Example 1; Page 98; 212pp; English.
 PS
 XX The invention relates to a DNA sequence which is a fully defined sequence
 CC of 103551 base pairs appearing as ABX04971, or its variant, that it is
 CC not a sequence encoding all or part amino acids 1-920 encoded by mon AI
 CC as given in the specification. The DNA is the S. cinamonensis polyketide
 CC antibiotic monensin biosynthetic gene cluster. Also included are a
 CC recombinant host cell which has been transformed to contain the gene
 CC cluster (and is capable of expressing a corresponding polypeptide), a
 CC hybridization probe derived from the gene cluster (for identification and
 CC isolation of the same or analogous gene cluster, e.g. one which binds
 CC specifically to a region of the monensin gene cluster selected from mon
 CC BI, mon BII, mon CI, mon CII, mon H, mon RI, mon RII, mon T, mon AIX and
 CC mon AX), the use of the mon RI gene or variant and a monensin promoter to
 CC control expression of a heterologous gene in Streptomyces cinamonensis,
 CC a polypeptide encoded by a portion of the monensin gene cluster
 CC (preferably comprising mon BI, mon BII, mon AIX or mon AX or their
 CC mutants, alleles or variants), an epoxidase enzyme encoded by mon CI, a
 CC cyclase enzyme encoded by mon CII, producing S. cinamonensis capable of
 CC enhanced levels of production of monensin comprising engineering it to
 CC overexpress the mon RI gene, S. cinamonensis containing multiple copies
 CC of the mon RI gene and/or its variants, expressing a gene heterologous to
 CC S. cinamonensis comprising transforming S. cinamonensis with DNA
 CC encoding a heterologous gene and expressing the gene under control of the
 CC activator gene mon RI or actII/orf4 and 13-propyl erythromycin A. The
 CC processes and materials (enzyme systems, nucleic acids and vectors) are
 CC useful for preparing polyketides by recombinant synthesis. The
 CC polyketides are useful as insecticides, antibiotics, antihelminthics,
 CC antifungals, antibacterials or other pharmaceuticals. In particular the
 CC gene is useful for the production of monensin, an antibiotic polyether
 CC polyketide. The present sequence represents a protein encoded by the
 CC monensin gene cluster
 XX
 XX Sequence 427 AA;
 SQ
 Query Match 62.2%; Score 46; DB 4; Length 427;
 Best Local Similarity 69.2%; Pred. No. 9;
 Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 FAVENPMNRLVAE 13
 ||: ||: |||

Db 161 FSVHEPMGRSLVAE 173

RESULT 7
ABB08908
ID ABB08908 standard; protein; 159 AA.
XX
AC ABB08908;
XX
DT 28-JUN-2002 (first entry)
XX
DE Pepper mild mottle virus envelope protein.
XX
KW Pepper mild mottle virus; PMV; envelope; Tobamovirus; transgenic plant;
KW plant expression vector; transgenic.
XX
OS Pepper mild mottle virus.
XX
FN KR149216-B1.
XX
PD 17-AUG-1998.
XX
PF 11-NOV-1994; 94KR-00029486.
XX
PR 11-NOV-1994; 94KR-00029486.
XX
PA (POHA-) POHANG ENG COLLEGE.
XX
PI Nam H, Choe J, Lee H, Park Y, Kim C;
XX
DR WPI; 2000-252824/22.
DR N-PSDB; ABA97721.
XX
PT ENVELOPE GENE OF PEPPER MILD MOTTLE VIRUS.
XX
PS Claim 2; Page 8; 12pp; Korean.
XX
CC The invention relates to the envelope gene (ABA97721) and protein (ABB08908) from pepper mild mottle virus (PMV - a member of the Tobamovirus family of single-stranded RNA viruses). The invention also encompasses plant expression vectors which comprise the PMV envelope gene. The present sequence represents the PMV envelope protein
SQ Sequence 159 AA;

Query Match 58.1%; Score 43; DB 3; Length 159;
Best Local Similarity 69.2%; Pred. No. 9.9;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 VENPMNRLVAETL 15
||| |
DB 97 VENPQNPTTAETL 109

RESULT 8
ABB71894
ID ABB71894 standard; protein; 307 AA.
XX
AC ABB71894;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 42474.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
FN WO200171042-A2.
XX
PD 27-SEP-2001.
XX

PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
XX
XX 11-JUL-2000; 2000US-00614150.
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABL15997.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions.
XX
PS Disclosure; SEQ ID NO 42474; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-ABBS72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 307 AA;

Query Match 55.4%; Score 41; DB 4; Length 307;
Best Local Similarity 64.3%; Pred. NO. 54;
Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 AVENPMNRLVAETL 15
||| |
DB 59 AVRNPRLTVEVL 72

RESULT 9
ADO08016
ID ADO08016 standard; protein; 307 AA.
XX
AC ADO08016;
XX
DT 01-JUL-2004 (first entry)
XX
DE Fly polypeptide #58.
XX
KW Fly; fat cell number; fat cell size; obesity; diabetes; anorectic;
KW antidiabetic.
XX
OS Diptera.
XX
FN US2004071700-A1.
XX
PD 15-APR-2004.
XX
PF 09-OCT-2002; 2002US-00267502.
XX
PR 09-OCT-2002; 2002US-00267502.
XX
PA (LIFE-) LIFE SCI DEV CORP.
XX
PI Kim J, Galant R;
XX
DR WPI; 2004-328526/30.
DR N-PSDB; ADO07799.
XX
PT Identifying compounds that influence fat cell number or size for treating or preventing obesity or diabetes by exposing the cell to the agent and identifying fat cell number or size relative to cells not exposed to the

PT agent.
 XX Claim 14; SEQ ID NO 342; 275pp; English.
 XX
 XX The invention relates to a method of identifying compounds that influence
 CC fat cell number or size comprising providing a cell that expresses a gene
 CC and an agent, exposing the cell to the agent and identifying fat cell
 CC number or size relative to cells not exposed to the agent. The method
 CC also comprises providing an expression vector and an agent, exposing the
 CC vector to the agent, detecting a change in expression of the gene
 CC relative to expression of the gene in an expression vector not exposed to
 CC the agent, treating a subject with the agent and identifying fat cell
 CC number or size in the subject. The agent comprises an antisense
 CC oligonucleotide. The subject comprises a mammal, preferably a human. The
 CC method also comprises providing a polypeptide and an agent, exposing the
 CC polypeptide to the agent, detecting binding of the agent to the
 CC polypeptide or a change in an activity of the polypeptide, treating a
 CC subject with the agent and identifying fat cell number or size in the
 CC subject. The agent comprises an antibody. A method of regulating fat cell
 CC number or size comprises providing a subject containing fat cells and an
 CC agent that changes the expression of a gene, and treating the subject
 CC with the agent under conditions so that fat cell size or number in the
 CC subject is altered. The method is useful for identifying compounds that
 CC influence fat cell number or size, for preparing a composition for
 CC treating or preventing obesity or diabetes. This sequence represents a
 CC fly polypeptide used in the scope of the invention.
 XX Sequence 307 AA;
 SQ
 Query Match 55.4%; Score 41; DB 8; Length 307;
 Best Local Similarity 64.3%; Pred. No. 54;
 Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 2 AVENPMNRLVAETL 15
 DB 59 AVRNPRLTVETL 72
 |||||
 |||||
 RESULT 10
 ADS27735
 ID ADS27735 standard; protein; 514 AA.
 XX
 AC ADS27735;
 XX
 DT 02-DEC-2004 (first entry)
 XX
 DE Bacterial polypeptide #16768.
 XX
 KW Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polypeptide.
 XX
 OS Bacteria.
 XX
 XX US2003233675-A1.
 XX
 XX 18-DEC-2003.
 XX
 XX 20-FEB-2003; 2003US-00369493.
 XX
 XX 21-FEB-2002; 2002US-0360039P.
 XX
 XX (CAOY/) CAO Y.
 XX (HINK/) HINKLE G J.
 XX (SLAT/) SLATER S C.
 XX (CHEN/) CHEN X.
 XX (GOLD/) GOLDMAN B S.
 XX
 PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX WPI; 2004-061375/06.
 DR
 XX New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 XX
 XX Claim 1; SEQ ID NO 16768; 122pp; English.
 XX
 CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition. Improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX
 XX Sequence 514 AA;
 SQ
 Query Match 55.4%; Score 41; DB 8; Length 514;
 Best Local Similarity 57.1%; Pred. No. 1e+02;
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 QY 2 AVENPMNRLVAETL 15
 DB 385 AEDNFINALLARTL 398
 |||||
 |||||
 RESULT 11
 ADU47425
 ID ADU47425 standard; protein; 59 AA.
 XX
 AC ADU47425;
 XX
 DT 10-FEB-2005 (first entry)
 XX
 DE Yarrowia lipolytica SEC65-related protein.
 XX
 KW Orotate-phosphoribosyl transferase; URA5 gene; SCS7 gene; SEC65 gene;
 KW secretory pathway; hybridisation technique; yeast.
 XX
 OS Yarrowia lipolytica.
 XX
 XX US2004229306-A1.
 XX
 XX 18-NOV-2004.
 XX
 XX 03-JUN-2003; 2003US-00454125.
 XX
 XX 16-MAY-2003; 2003US-0471435P.
 XX
 XX (NETT/) NETT J H.
 XX
 XX Nett JH;
 XX
 XX WPI; 2004-813249/80.
 DR

XX New isolated orotate-phosphoribosyl transferase polynucleotides and
PT polypeptides, useful for producing and selecting yeast strains capable of
PT stable genetic integration of heterologous sequences into a host genome.
XX
XX Example 1; SEQ ID NO 20; 38pp; English.
XX
XX The present invention provides polypeptides encoding a novel orotate-
CC phosphoribosyl transferase (URA5) gene. The invention is useful for
CC producing and selecting yeast strains capable of stable genetic
CC integration of heterologous sequences into a host genome. The invention
CC is also useful for measuring expression of the SC97 gene and secretory
CC pathway (SEC65) gene and to further characterise the structure and
CC function of this gene and its encoded protein. The nucleic acid fragments
CC of the invention can also be used as probes for various hybridisation
CC techniques. The present sequence is Yarrowia lipolytica SEC65-related
CC protein.
XX
XX Sequence 59 AA;
SQ
Query Match 54.1%; Score 40; DB 8; Length 59;
Best Local Similarity 58.3%; Pred. No. 11;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
OY 2 AVENPMNRLVAE 13
Db |||||: :|||
22 AVENPLGQTIAE 33
RESULT 12
ID ABB69884 standard; protein; 518 AA.
XX
AC ABB69884;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 36444.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
XX
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
DR N-PSDB; ABL13987.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
XX Disclosure; SEQ ID NO 36444; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-

CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 518 AA;
SQ
Query Match 52.7%; Score 39; DB 4; Length 518;
Best Local Similarity 60.0%; Pred. No. 2.5e+02;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
OY 1 FAVENPMNRLVAETL 15
Db |||||: :|||
47 FRVERDQYKLVAE 61
RESULT 13
ID ADF06956 standard; protein; 559 AA.
XX
AC ADF06956;
XX
DT 12-FEB-2004 (first entry)
XX
DE Bacterial polypeptide #3069.
XX
KW Proteus mirabilis infection; bacterial infection; antibacterial;
KW immunostimulant.
XX
OS Proteus mirabilis.
XX
XX US6605709-B1.
XX
PD 12-AUG-2003.
XX
PF 05-APR-2000; 2000US-00543681.
XX
PR 09-APR-1999; 99US-0128706P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Breton GL;
XX
DR WPI; 2003-895291/82.
DR N-PSDB; ADF02784.
XX
XX New Proteus mirabilis polypeptides and polynucleotides, useful as
PT reagents for diagnosis of bacterial disease, as components of
PT antibacterial vaccines, as targets for antibacterial drugs, or as
PT biocontrol agents for plants.
XX
XX Disclosure; SEQ ID NO 7241; 870pp; English.
XX
XX The invention relates to new Proteus mirabilis polypeptides and
CC polynucleotides. The invention also relates to antibodies against the
CC polypeptides, methods for producing the polypeptides, a method of
CC generating vaccines for immunising an individual against P. mirabilis, a
CC method for evaluating a compound for the ability to bind a P. mirabilis
CC polypeptide and a method for screening test compounds for anti-bacterial
CC activity. The polypeptides and polynucleotides are useful as molecular
CC targets for diagnosing, preventing and treating pathological conditions
CC resulting from bacterial infection, as reagents for diagnosis of
CC bacterial diseases, as components of antibacterial vaccines, as targets
CC for antibacterial drugs or as bio-control agents for plants. This
CC sequence represents a Proteus mirabilis polypeptide of the invention.
XX
XX Sequence 559 AA;
SQ
Query Match 52.7%; Score 39; DB 7; Length 559;
Best Local Similarity 42.9%; Pred. No. 2.7e+02;
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
OY 1 FAVENPMNRLVAET 14
||:| |::| |

Db 397 PSLEKPVNRILINT 410

RESULT 14
ADJ49966

ID ADJ49966 standard; protein; 756 AA.

XX AC ADJ49966;

XX DT 06-MAY-2004 (first entry)

XX DE Oil-associated gene related protein #1466.

XX KW oil-associated gene; transgenic; enhanced seed oil; vegetable oil.

XX OS Unidentified.

XX PN US2004025202-A1.

XX PD 05-FEB-2004.

XX PF 14-MAR-2003; 2003US-00389566.

XX PR 15-MAR-2002; 2002US-0365301P.

XX PR 26-JUN-2002; 2002US-0391786P.

XX PR 26-JUN-2002; 2002US-0392018P.

XX PA (LAUR/) LAURIE C C.

XX PA (RAVA/) RAVANELLO M.

XX PA (SAVA/) SAVAGE T.

XX PA (LEDE/) LEDEAUX J R.

XX PA (ROGE/) ROGERS J A.

XX PI Laurie CC, Ravanello M, Savage T, Ledeaux JR, Rogers JA;

XX WPI; 2004-142683/14.

XX DR Novel recombinant DNA construct comprising a promoter functional in

PT plants operably linked to an oil-associated gene for producing transgenic

PT plant seed.

XX PS Example 3; SEQ ID NO 1970; 22pp; English.

XX CC The invention relates to a recombinant DNA construct comprising a

CC promoter functional in plants operably linked to an oil-associated gene.

CC The construct is useful for transgenic plant seed which has in its genome

CC the construct, that is functional in the plant to transcribe the oil-

CC associated gene. The transgenic plant seed grows into a plant having

CC enhanced seed oil as compared to wild type. The construct is useful for

CC producing hybrid maize seed. The transgenic plant seed is useful for

CC producing vegetable oil. The present sequence represents the amino acid

CC sequence of an oil-associated gene related protein.

XX SQ Sequence 756 AA;

Query Match 52.7%; Score 39; DB 8; Length 756;

Best Local Similarity 63.8%; Pred. No. 3.9e+02;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 FAVENPMNRILV 11

Db 67 FAKENPMNRVL 77

RESULT 15
AAU34800

ID AAU34800 standard; protein; 804 AA.

XX AC AAU34800;

XX DT 14-FEB-2002 (first entry)

XX DE E. coli cellular proliferation protein #381.

XX KW Antisense; prokaryotic cellular proliferation protein; antibiotic;

XX KW antibacterial; drug design.

XX OS Escherichia coli.

XX PN WO200170955-A2.

XX PD 27-SEP-2001.

XX PF 21-MAR-2001; 2001WO-US009180.

XX PR 21-MAR-2000; 2000US-0191078P.

XX PR 23-MAY-2000; 2000US-0206848P.

XX PR 26-MAY-2000; 2000US-020727P.

XX PR 23-OCT-2000; 2000US-0242578P.

XX PR 27-NOV-2000; 2000US-0253625P.

XX PR 22-DEC-2000; 2000US-0257931P.

XX PR 16-FEB-2001; 2001US-0269308P.

XX PA (ELIT-) ELITRA PHARM INC.

XX PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

XX PI Yamamoto RT, Xu HH;

XX WPI; 2001-611495/70.

XX DR N-PSDB; AAS52659.

XX PT New polynucleotides for the identification and development of

PT antibiotics, comprise sequences of antisense nucleic acids.

XX PS Example 3; SEQ ID NO 10393; 511pp; English.

XX CC The invention relates to antisense inhibitors of genes essential to

CC prokaryotic cellular proliferation, their use in identifying the genes,

CC their use in the discovery of novel antibiotics, the essential genes

CC themselves and the encoded proteins. The prokaryotes used are Escherichia

CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,

CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also

CC useful for the identification of potential new targets for antibiotic

CC development. The antisense nucleic acids can also be used to identify

CC proteins used in proliferation, to express these proteins, and to obtain

CC antibodies capable of binding to the expressed proteins. The proteins can

CC be used to screen compounds in rational drug discovery programmes. The

CC antisense nucleic acid sequence is also useful to screen for homologous

CC nucleic acids which are required for cell proliferation in a wide variety

CC of organisms. The present sequence represents an essential prokaryotic

CC cellular proliferation protein. Note: The sequence data for this patent

CC did not form part of the printed specification, but was obtained in

CC electronic format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 804 AA;

Query Match 52.7%; Score 39; DB 4; Length 804;

Best Local Similarity 64.3%; Pred. No. 4.2e+02;

Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 AVENPMNRILVAETL 15

Db 348 AVEQQMNELLAEVL 361

RESULT 16
ABU28825

ID ABU28825 standard; protein; 804 AA.

XX AC ABU28825;

XX DT 19-JUN-2003 (first entry)

XX DE Protein encoded by Prokaryotic essential gene #14352.

KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX Escherichia coli.
 OS
 XX WO200277183-A2.
 PN
 XX 03-OCT-2002.
 PD
 XX
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 PF 21-MAR-2001; 2001US-00815242.
 PR
 PR 06-SEP-2001; 2001US-00948993.
 PR
 PR 25-OCT-2001; 2001US-0342923P.
 PR
 PR 08-FEB-2002; 2002US-00072851.
 PR
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 XX (ELIT-) ELITRA PHARM INC.
 PA
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KU, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 PI
 DR WPI; 2003-029926/02.
 XX
 DR N-PSDB; ACA32695.
 XX
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 56749; 1766pp; English.
 XX
 XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 804 AA;

 Query Match 52.7%; Score 39; DB 6; Length 804;
 Best Local Similarity 64.3%; Pred. No. 4.2e+02;
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0

 Qy 2 AVENPMRRLVAETL 15
 Db 348 AVEQQMNELLAEYL 361

RESULT 17
ADN25977
ID ADN25977 standard; protein; 711 AA.
XX
XX ADN25977;
AC
XX
XX 02-DEC-2004 (first entry)
DT
XX
XX Bacterial polypeptide #8630.
DE
XX
XX Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX
XX
XX Bacteria.
OS
XX
XX US2003233675-A1.
PN
XX
XX 18-DEC-2003.
PD
XX
XX 20-FEB-2003; 2003US-00369493.
PF
XX
XX 21-FEB-2002; 2002US-0360039P.
PR
XX
XX (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
XX
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
PI WPI; 2004-061375/06.
PI
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
XX
XX Claim 1; SEQ ID NO 8630; 122pp; English.
PS
XX
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
XX Sequence 711 AA;
SQ

Matches 9; Conservative 3; Mismatches 2; Indels 1; Gaps 1;
QY 2 AVENPMNR-LVAETL 15
| :||:| | :||:| |
Db 605 AEDNPLNRMLIAEQL 619

RESULT 18
ABP06885
ID ABP06885 standard; protein; 59 AA.
XX ABP06885;
XX
DT 24-JUN-2002 (first entry)
XX
DE Human ORFX protein sequence SEQ ID NO:13752.
XX
KW Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
KW hypertension; hypothyroidism; cholesterol ester storage disease;
KW immune deficiency; immune disorder; infectious disease;
KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
KW myaesthesia gravis.
XX
OS Homo sapiens.
XX
PN WO200192523-A2.
XX
PD 06-DEC-2001.
XX
PF 29-MAY-2001; 2001WO-US010836.
XX
PR 30-MAY-2000; 2000US-0206132P.
PR 29-AUG-2000; 2000US-0228716P.
XX
XX (CURA-) CURAGEN CORP.
PA
XX
XX Shinkets RA, Leach WD;
XX
XX WPI; 2002-106308/14.
DR N-PSDB; ABN22637.
XX
XX Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and autoimmune disorders.
XX
XX Disclosure; SEQ ID NO 13752; 1037pp; English.
XX
XX The present invention describes substantially purified human proteins
CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
CC in the specification). ABN15762 to ABN27252 encode the human ORFX
CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
CC treating or preventing a pathology associated with an ORFX-associated
CC disorder in humans, and in the manufacture of a medicament for treating a
CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
CC sequences can be used in gene therapy. ORFX sequences can be used in the
CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
CC osteoarthritis, neurodegenerative disorders, disorders related to organ
CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
CC storage disease, various immune deficiencies and disorders, infectious
CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
CC arthritis, autoimmune thyroiditis, myaesthesia gravis, graft-versus-host
CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
CC bone degenerative disorders, or periodontal disease, and for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues and conditions resulting from
CC systemic cytokine damage. N.B. The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic

CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 59 AA;
Query Match 51.4%; Score 38; DB 5; Length 59;
Best Local Similarity 50.0%; Pred. No. 26;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
QY 3 VENPMNRLVAET 14
: ||:| | :||:| |
Db 42 IRNPVSRLIATT 53

RESULT 19
ABU18938
ID ABU18938 standard; protein; 252 AA.
XX ABU18938;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #4465.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Bacillus anthracis.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Wang L, Zamudio C, Malone C, Haseibeck R, Ohlsen KL, Zyskind JW;
PI Walli D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
XX WPI; 2003-029926/02.
DR N-PSDB; ACA22808.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 25; SEQ ID NO 46862; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway;
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of

DT 28-MAR-2001 (first entry)
 XX Arabidopsis acyl-CoA oxidase ATACX1.
 DE Acyl-CoA oxidase; ACOX; ATACX1; fatty acid; beta-oxidation;
 KW transgenic plant; lipid; vegetable oil.
 KW Arabidopsis thaliana.
 OS Arabidopsis thaliana.
 XX Key Location/Qualifiers
 FH Misc-difference 519
 FT /note= "encoded by CCT"
 XX WO200075350-A2.
 XX 14-DEC-2000.
 XX 08-JUN-2000; 2000WO-US016149.
 XX 08-JUN-1999; 99US-0138162P.
 XX (CALJ) CALGENE LLC.
 XX Dehesh K;
 XX WPI: 2001-091217/10.
 XX N-PSDB; AAA89274.
 XX Polynucleotide encoding proteins involved in fatty acid beta-oxidation,
 PT useful for modifying oil or lipid composition in host plant cells.
 XX Claim 27; Page 33-35; 59pp; English.
 XX The present sequence is that of ATACX1, an acyl-CoA oxidase (ACOX) of
 CC Arabidopsis thaliana. ACOX enzymes catalyzes the initial step of fatty
 CC acid beta-oxidation. The invention provides ACOX and 3-ketoacyl-CoA
 CC thiolase (thiolase) isolated DNA sequences (see AAA89273-85) and encoded
 CC polypeptides (see AAB19999-B20011), and methods of producing such
 CC polypeptides using recombinant methods in host (bacterial, insect,
 CC mammalian and especially plant) cells. Expression of ACOX and/or thiolase
 CC in a plant can be increased or suppressed using nucleic acid constructs
 CC in sense or antisense orientation. This provides a means of manipulating
 CC the content and composition of fatty acids and compounds containing such
 CC fatty acids, such as oils, waxes, fats and storage proteins. Suppression
 CC of ACOX and/or thiolase expression also provides a means of inhibiting or
 CC postponing seed germination
 XX Sequence 664 AA;
 SQ Query Match 51.4%; Score 38; DB 4; Length 664;
 Best Local Similarity 60.0%; Pred. No. 5.2e+02;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 QY 1 FAVENPMNRLVAETL 15
 Db 32 FDVSNRMSRLVANDL 46
 RESULT 23
 ABO77366
 ID ABO77366 standard; protein; 860 AA.
 XX ABO77366;
 AC ABO77366;
 XX 29-JUL-2004 (first entry)
 DT Pseudomonas aeruginosa polypeptide #9541.
 XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
 DE Pseudomonas aeruginosa.
 KW Pseudomonas aeruginosa.
 OS US6551795-B1.

XX 22-APR-2003.
 PD 18-FEB-1999; 99US-00252991.
 PF 18-FEB-1998; 98US-0074788P.
 PR 27-JUL-1998; 98US-0094190P.
 XX (GENO-) GENOME THERAPEUTICS CORP.
 XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
 PI WPI: 2003-615309/58.
 XX N-PSDB; ABD10937.
 DR Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide, of
 PT useful as molecular targets for diagnostics, prophylaxis and treatment of
 PT pathological conditions resulting from bacterial infection.
 XX Disclosure; SEQ ID NO 26112; 455pp; English.
 XX The invention relates to Pseudomonas aeruginosa polypeptides and the
 CC polynucleotides encoding them. The sequences are useful in diagnosis and
 CC therapy of pathological conditions, as molecular targets for diagnostics,
 CC prophylaxis and treatment of pathological conditions resulting from a
 CC bacterial infection, for evaluating a compound, such as a polypeptide,
 CC for the ability to bind a P. aeruginosa nucleic acid, as components of
 CC effective antibacterial targets, as targets for antibacterial drugs,
 CC including anti-P. aeruginosa drugs, as templates for recombinant
 CC production of P. aeruginosa-derived peptides or polypeptides, as target
 CC components for diagnosis and/or treatment of P. aeruginosa-caused
 CC infection, and in detection of P. aeruginosa sequences or other sequences
 CC of Pseudomonas species using biochip technology. Sequences ABO67826-
 CC ABO84336 represent P. aeruginosa polypeptides of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format from USPTO at
 CC seqdata.uspto.gov/sequence.html
 XX Sequence 860 AA;
 SQ Query Match 51.4%; Score 38; DB 7; Length 860;
 Best Local Similarity 58.3%; Pred. No. 7.2e+02;
 Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
 QY 4 ENPMNRLVAETL 15
 Db 744 DNPVNQLVAKGL 755
 RESULT 24
 ABB53002
 ID ABB53002 standard; protein; 956 AA.
 XX ABB53002;
 AC ABB53002;
 XX 11-FEB-2002 (first entry)
 DT Escherichia coli polypeptide SEQ ID NO 1353.
 XX Escherichia coli; B2/D+A-; antiinflammatory; antibacterial;
 KW immunosuppressive; extra-intestinal infection; phylogeny; meningitis;
 KW systemic infection; non-diarrhoeal infection; septicemia;
 KW pyelonephritis; antibiotic resistance.
 XX Escherichia coli.
 OS WO200166572-A2.
 XX 13-SEP-2001.
 PD 12-MAR-2001; 2001WO-BF003445.
 PF 10-MAR-2000; 2000FR-00003145.
 XX

CC polynucleotide-packaging polypeptide comprising high mobility group (HMG)
 CC box domains and a protein transduction domain operably linked to a
 CC targeting signal for directing the recombinant polynucleotide-packaging
 CC polypeptide to an intracellular organelle, where the polypeptide
 CC associates with about 10-100 nucleotides of the polynucleotide and
 CC packages the polynucleotide for delivery to the targeted organelle. The
 CC invention also relates to a recombinant high mobility protein or its
 CC fragment comprising a protein transduction domain operably linked to a
 CC targeting signal for a non-nuclear organelle, where the recombinant high
 CC mobility protein associates with a polynucleotide under low stringency
 CC conditions and packages the polynucleotide for delivery to the non-
 CC nuclear organelle, a method of packaging a polynucleotide involving
 CC combining a polynucleotide with an amount of a recombinant polypeptide
 CC sufficient to package the polynucleotide, where the recombinant
 CC polypeptide comprises two or more HMG box domains and a protein
 CC transduction domain operably linked to an organelle targeting signal, a
 CC method of delivering a polynucleotide to an organelle of a cell and a
 CC method of treating a host involving contacting at least one of the host's
 CC cells with a complex comprising an amount of a recombinant polypeptide
 CC effective to package a polynucleotide encoding a therapeutic polypeptide.
 CC The compositions are useful for treating gene related diseases or
 CC disorders chosen from cancer, Alzheimers disease, Parkinsons disease,
 CC hypercholesterolemia, cystic fibrosis, anemia, diabetes, arthritis and
 CC autoimmune diseases, and for treating diseases caused by mitochondrial
 CC genetic defects or abnormalities such as Alpers disease, beta-oxidation
 CC defects, carnitine-acyl-carnitine deficiency, mitochondrial cytopathy and
 CC myogastrintestinal encephalomyopathy. This sequence represents a
 CC mitochondrial localization signal used in the scope of the invention.

XX Sequence 324 AA;

Query Match 50.7%; Score 37.5; DB 9; Length 324;

Best Local Similarity 45.0%; Pred. No. 2.7e+02;

Matches 9; Conservative 2; Mismatches 2; Indels 7; Gaps 1;

QY 1 FAVENP-----MNRVAV 13

Db 290 FTAENPLHQPSPSLNKLVAE 309

RESULT 27

ABP02493

ID ABP02493 standard; protein; 88 AA.

XX AC ABP02493;

XX 25-JUN-2002 (first entry)

DE Human ORFX protein sequence SEQ ID NO:4968.

XX Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
 KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
 KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
 KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
 KW hypertension; hypothyroidism; cholesterol ester storage disease;
 KW immune deficiency; immune disorder; infectious disease;
 KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
 KW myasthenia gravis.

OS Homo sapiens.

XX WO200192523-A2.

PN 06-DEC-2001.

PD 29-MAY-2001; 2001WO-US010836.

PF 30-MAY-2000; 2000US-0206132P.

PR 29-AUG-2000; 2000US-0228716P.

XX (CURA-) CURAGEN CORP.

PA Shimketa RA, Leach MD;

XX PI

XX

DR WPI; 2002-106308/14.

DR N-PSDB; ABN18245.

XX

PT Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and autoimmune disorders.

XX

PS Disclosure; SEQ ID NO 4968; 1037pp; English.

XX

CC The present invention describes substantially purified human proteins
 CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
 CC in the specification). ABN15762 to ABN27252 encode the human ORFX
 CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
 CC treating or preventing a pathology associated with an ORFX-associated
 CC disorder in humans, and in the manufacture of a medicament for treating a
 CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
 CC sequences can be used in gene therapy. ORFX sequences can be used in the
 CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
 CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
 CC osteoarthritis, neurodegenerative disorders, disorders related to organ
 CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
 CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
 CC storage disease, various immune deficiencies and disorders, infectious
 CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
 CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
 CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
 CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
 CC bone degenerative disorders, or periodontal disease, and for gut
 CC protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues and conditions resulting from
 CC systemic cytokine damage. N.B. The sequence data for this patent did not
 CC form part of the printed specification, but was obtained in electronic
 CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 88 AA;

Query Match 50.0%; Score 37; DB 5; Length 88;

Best Local Similarity 50.0%; Pred. No. 66;

Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 4 ENPMNRLVAETL 15

Db 12 DNPVNRVTVEAM 23

RESULT 28

AAW05273

ID AAW05273 standard; protein; 115 AA.

XX AC AAW05273;

XX 25-MAR-2003 (revised)

DT 12-NOV-1996 (first entry)

XX

DE N-terminally truncated T cell replacing factor (19-133).

XX

KW BCGF II; B cell growth factor; T cell replacing factor; TCR;
 KW B cell differentiation factor; N-terminal truncation; mutant; variant;
 KW modification; cancer; autoimmune disease; infection; treatment;
 KW diagnosis; prevention.

XX

OS Mammalia.

XX

XX Key Location/Qualifiers

FT Protein 1..115

FT /note "corresponds to TCR protein from which the first

FT 18 N-terminal amino acids have been deleted"

XX

PN JP08067698-A.

XX

XX 12-MAR-1996.

XX

```

PF 04-JUL-1986; 95JP-00195345.
XX
PR 04-JUL-1986; 86JP-00157227.
XX
XX (MOTO/) MOTOCHIKA T.
PA
XX WPI; 1996-196574/20.
DR
DR N-PSDB; AAT14922.
XX
PT B cell differentiation factor designated T cell replacing factor (TRF) -
PT useful for prevention, diagnosis and treatment of infectious and
PT auto-immune diseases and cancer.
XX
PS Claim 4; Page 2-3; lipp; Japanese.
XX
CC The present sequence is that of an N-terminally truncated form of a
CC protein designated T cell replacing factor (TCR) which has B cell growth
CC factor (BCGF II) activity. The truncated protein, deleted up to and
CC including amino acid residue Ala18, can be used, in common with the full-
CC length protein for prevention, diagnosis and treatment of infectious
CC diseases, autoimmune diseases and certain cancers. (Updated on 25-MAR-
CC 2003 to correct PF field.)
XX
XX Sequence 115 AA;
SQ
Query Match 50.0%; Score 37; DB 2; Length 115;
Best Local Similarity 57.1%; Pred. No. 92;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 AVENPMNRLVAETL 15
DB |:| | | | | | |
2 AWEIPMSTVVKETL 15

RESULT 29
AAW72948
ID AAW72948 standard; protein; 115 AA.
XX
AC AAW72948;
XX
DT 21-JAN-1999 (first entry)
XX
DE Truncated T cell replacing factor.
XX
DE Mouse; T cell replacing factor; TRF; B cell growth factor; BCGF;
KW B cell differentiation activating factor; B cell differentiating factor.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT Misc-difference 1 /label= optional
FT Misc-difference 2 /label= optional
FT Misc-difference 3 /label= optional
FT
FT
XX JP10262673-A.
XX
XX 06-OCT-1998.
PD
XX
XX 04-JUL-1986; 97JP-00055274.
PF
XX
XX 04-JUL-1986; 95JP-00195345.
PR
XX
XX (HONJ/) HONJO T.
PA
XX WPI; 1998-587290/50.
DR
XX
PT Preparation of B cell differentiating factor - by culturing transformed
PT cells containing plasmids with DNA encoding the factor.
XX
XX Claim 6; Page 3; 10pp; Japanese.
PS

XX The present sequence represents a truncated T cell replacing factor (full
CC length given in AAW72947) having B cell growth factor activity. A method
CC has been developed for the preparation of B cell differentiation
CC activating factor (also designated BCGF. B cell growth factor, T cell
CC replacing factor and TRF). The method comprises producing a protein by
CC culturing a transformed cell introduced by a plasmid containing a DNA
CC coding B cell differentiation activating factor. The method is used for
CC the large scale preparation of B cell differentiating factor
XX
XX Sequence 115 AA;
SQ
Query Match 50.0%; Score 37; DB 2; Length 115;
Best Local Similarity 57.1%; Pred. No. 92;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 AVENPMNRLVAETL 15
DB |:| | | | | | |
2 AWEIPMSTVVKETL 15

RESULT 30
AAP71064
ID AAP71064 standard; protein; 133 AA.
XX
AC AAP71064;
XX
DT 25-MAR-2003 (revised)
DT 13-MAY-1991 (first entry)
XX
DE Murine eosinophil differentiation factor (EDF).
XX
XX EDA; B-cell growth factor II; BCGF II; haemopoiesis; myeloma cells;
KW eosinophilia; interleukin.
XX
OS Mus musculus.
XX
XX W08704466-A.
XX
XX 30-JUL-1987.
PD
XX
XX 15-JAN-1987; 87WO-GB0000021.
PF
XX
XX 15-JAN-1986; 86GB-00000835.
PR
XX 15-JAN-1986; 86GB-00000836.
XX
XX (AMSH ) AMERSHAM INT PLC.
PA
XX
XX Sanderson CJ, Young IG;
PI
XX WPI; 1987-221267/31.
DR
XX
XX Eosinophil differentiation factor - with eosinophil differentiation
PT activity and B cell growth promoting activity of B cell growth factor II.
XX
XX Claim 7; Fig 5; 37pp; English.
PS
XX
XX The EDF gene product is a novel interleukin, useful in the study of
CC haemopoiesis and B-cell differentiation, and may have utility in MAb
CC production. It may be used therapeutically in regulation of the immune
CC response, and promotion of eosinophilia. See also AAN71243. (Updated on
CC 25-MAR-2003 to correct PA field.)
XX
XX Sequence 133 AA;
SQ
Query Match 50.0%; Score 37; DB 1; Length 133;
Best Local Similarity 57.1%; Pred. No. 1.1e+02;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 AVENPMNRLVAETL 15
DB |:| | | | | | |
20 AWEIPMSTVVKETL 33

```

Search completed: December 17, 2005, 07:29:05
Job time : 194 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2005, 07:21:37 ; Search time 38 Seconds
(without alignments)
37.980 Million cell updates/sec

Title: US-10-787-382-20

Perfect score: 74

Sequence: 1 PAVENPMRLVAETL 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : PIR 80.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	74.3	134	2 JCS116	interleukin-5 prec
2	43	58.1	157	1 VCTMPV	coat protein - pep
3	41	55.4	514	2 A87266	sensor histidine k
4	40	54.1	310	2 JCS516	signal recognition
5	40	54.1	447	1 E69878	RNA-binding Sun pr
6	40	54.1	1944	2 AH3098	rhizobiotin/RTX to
7	40	54.1	1990	2 A96188	probable phosphos
8	39	52.7	158	2 A60023	coat protein - odo
9	39	52.7	158	2 S14468	coat protein - odo
10	39	52.7	158	2 JQ1266	coat protein - odo
11	39	52.7	221	2 T02666	germin-like protei
12	39	52.7	756	2 JCS590	cycloartenol synth
13	39	52.7	804	1 ISECTB	DNA topoisomerase
14	39	52.7	804	2 B91208	hypothetical prote
15	39	52.7	804	2 D86054	germin-like protei
16	38	51.4	181	2 T02591	probable rna matur
17	38	51.4	773	2 T04094	probable sensor/re
18	38	51.4	786	2 F83292	hypothetical prote
19	37	50.0	50	2 H97798	hypothetical prote
20	37	50.0	132	1 A48418	interleukin-5 prec
21	37	50.0	133	1 ICMS5	interleukin-5 prec
22	37	50.0	159	4 S34856	coat protein mutan
23	37	50.0	162	1 D71206	hypothetical prote
24	37	50.0	162	2 H75019	hypothetical prote
25	37	50.0	311	2 E85565	probable tRNA synt
26	37	50.0	311	2 B90715	probable tRNA synt
27	37	50.0	311	2 A64800	purine nucleosidas
28	37	50.0	493	2 D42519	Al8R protein - vac
29	37	50.0	557	2 S23429	protein kinase ERK

30	37	50.0	650	2 T41681	probable serine/th
31	37	50.0	943	2 T33795	hypothetical prote
32	37	50.0	1304	2 T14073	dynein lb heavy ch
33	37	50.0	1384	2 T18366	latrophilin-2, epl
34	37	50.0	1397	2 T18377	latrophilin-2, epl
35	37	50.0	1399	2 T18370	latrophilin-2, epl
36	37	50.0	1412	2 T18380	latrophilin-2, epl
37	37	50.0	1450	2 T18382	latrophilin-2, epl
38	37	50.0	1452	2 T17157	Cu2AA protein - ra
39	37	50.0	1463	2 T18386	latrophilin-2, epl
40	37	50.0	1463	2 T17159	Cu2AC protein - ra
41	37	50.0	1465	2 T18384	latrophilin-2, epl
42	37	50.0	1467	2 T17160	Cu2BA protein - ra
43	37	50.0	1478	2 T18388	latrophilin-2, epl
44	37	50.0	1478	2 T17185	CL2BC protein - ra
45	37	50.0	1487	2 T14324	alpha-latrotoxin r
46	37	50.0	2510	2 T28160	hypothetical prote
47	36	48.6	157	2 AH3460	arsenate reductase
48	36	48.6	166	2 T42290	hypothetical prote
49	36	48.6	200	2 G82139	endopeptidase Clp
50	36	48.6	201	1 R5EB4Y	ribosomal protein
51	36	48.6	201	1 R5EC4	ribosomal protein
52	36	48.6	201	2 AB1006	50S ribosomal chai
53	36	48.6	201	2 H91151	ribosomal protein
54	36	48.6	201	2 D85937	ribosomal protein
55	36	48.6	201	2 AC0026	50S ribosomal prot
56	36	48.6	247	2 H69374	NH(3)-dependent NA
57	36	48.6	251	2 H72275	phosphate ABC tran
58	36	48.6	255	2 A81099	1-acyl-sn-glycerol
59	36	48.6	255	2 A81842	1-acylglycerol-3-p
60	36	48.6	255	2 S70545	1-acylglycerol-3-p
61	36	48.6	266	2 S67435	phosphoadenylyl-su
62	36	48.6	360	2 AF0247	probable exported
63	36	48.6	387	2 H88941	protein R13D11.4 l
64	36	48.6	394	2 H87640	sensor histidine k
65	36	48.6	486	2 A1575	acetaldehyde dehyd
66	36	48.6	486	2 AC1222	acetaldehyde dehyd
67	36	48.6	513	2 H69735	endo-1,4-beta-xyla
68	36	48.6	596	2 A11865	hypothetical prote
69	36	48.6	615	2 D87344	60 kd inner-membra
70	36	48.6	651	2 C75014	hypothetical prote
71	36	48.6	654	2 F85057	receptor-like prot
72	36	48.6	804	2 AD0958	DNA Gyrase Chain B
73	36	48.6	1089	2 T31583	hypothetical prote
74	36	48.6	1140	2 T41457	dna repair protein
75	36	48.6	1447	2 T15200	hypothetical prote
76	36	48.6	3788	2 T30851	lysosomal traffick
77	36	48.6	3788	2 T13960	beige protein homo
78	35	47.3	139	2 E72559	hypothetical prote
79	35	47.3	163	2 E90212	conserved hypotet
80	35	47.3	185	2 C95109	ribosome recycling
81	35	47.3	185	2 F97977	ribosome recycling
82	35	47.3	227	2 S52856	AraC-like protein
83	35	47.3	251	2 AE2749	transcription regu
84	35	47.3	254	2 E84528	urase accessory p
85	35	47.3	256	2 D71979	signal recognition
86	35	47.3	273	2 S21731	probable hydroxybu
87	35	47.3	276	2 C95891	transcription regu
88	35	47.3	278	2 D97530	hypothetical prote
89	35	47.3	286	2 G95301	conserved hypotet
90	35	47.3	288	2 D69480	two-component sens
91	35	47.3	334	2 H84138	cytochrome d ubiq
92	35	47.3	378	2 A82406	ribosomal protein
93	35	47.3	390	2 S54026	branched chain ani
94	35	47.3	405	2 T50717	probable O-antigen
95	35	47.3	410	2 AH0051	gamma-glutamyl pho
96	35	47.3	416	2 AH0543	glutamate-5-semial
97	35	47.3	417	1 RDECER	gamma-glutamylphos
98	35	47.3	417	2 D85513	gamma-glutamylphos
99	35	47.3	417	2 F90662	probable fibrial
100	35	47.3	441	2 AC0544	conserved domain p
101	35	47.3	442	2 B55145	hypothetical prote
102	35	47.3	447	2 H98012	

103 35 47.3 449 2 JC1137
104 35 47.3 449 2 S18561
105 35 47.3 479 2 G75092
106 35 47.3 484 2 AE3176
107 35 47.3 493 2 C36415
108 35 47.3 493 2 T37405
109 35 47.3 493 2 A36850
110 35 47.3 493 2 T28560
111 35 47.3 493 2 H72165
112 35 47.3 496 2 AD3085
113 35 47.3 496 2 E98201
114 35 47.3 499 2 B75577
115 35 47.3 513 2 A33966
116 35 47.3 532 2 S40983
117 35 47.3 537 2 AF0575
118 35 47.3 544 2 S61302
119 35 47.3 591 2 S73790
120 35 47.3 649 2 T47609
121 35 47.3 664 2 G84771
122 35 47.3 682 2 B86336
123 35 47.3 747 2 F88561
124 35 47.3 759 2 E81869
125 35 47.3 886 2 B83306
126 35 47.3 1040 2 AH1926
127 35 47.3 1377 2 T19214
128 34.5 46.6 256 2 B93383
129 34.5 46.6 327 2 T04004
130 34.5 46.6 427 2 AE2851
131 34.5 46.6 427 2 C97628
132 34 45.9 90 2 H83931
133 34 45.9 111 2 S75168
134 34 45.9 144 2 T51597
135 34 45.9 154 2 AB2509
136 34 45.9 158 1 VCTMKO
137 34 45.9 158 2 S73790
138 34 45.9 158 2 S54374
139 34 45.9 158 2 S55374
140 34 45.9 159 1 VCTMVU
141 34 45.9 159 2 D81158
142 34 45.9 159 2 H81946
143 34 45.9 166 1 B71033
144 34 45.9 166 2 A75181
145 34 45.9 219 1 A39574
146 34 45.9 241 1 JQ1641
147 34 45.9 257 2 S70544
148 34 45.9 264 2 C84971
149 34 45.9 267 2 AG0208
150 34 45.9 271 2 S18730

ALIGNMENTS

RESULT 1
JC5116
interleukin-5 precursor - bovine
C:Species: Bos primigenius taurus (cattle)
C>Date: 02-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 09-Jul-2004
C:Accession: JC5116
R/Mertens, B.; Gobright, E.; Seow, H.F.
Gene 176, 273-274, 1996
A/Title: The nucleotide sequence of the bovine interleukin-5 encoding cDNA.
A/Reference number: JC5116; MUID:97075944; PMID:8918267
A/Molecule type: mRNA
A/Residues: 1-134 <MER>
A/Cross-references: UNIPROT:P52173; UNIPARC:UPI000012D4CB; EMBL:Z67872; NID:g1113120; PI
A/Experimental source: lymphocytes
C/Comment: This protein plays a role in the eosinophilia associated with parasitic disea
C/Genetics:
C/Superfamily: interleukin-5
C/Keywords: cytokine; glycoprotein; growth factor; homodimer; lymphokine; T-cell

F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-134/Product: interleukin-5 #status predicted <MAT>
F:63/Disulfide bonds: carboxylate (to 105) #status predicted
F:76-90/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:105/Disulfide bonds: interchain (to 63) #status predicted
Query Match 74.3%; Score 55; DB 2; Length 134;
Best Local Similarity 85.7%; Pred. No. 0.015;
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 2 AVENPMNRLVAETL 15
DB 21 AVESMTNRLVAETL 34
||||:|||||
RESULT 2
VCTMPV
coat protein - pepper mild mottle virus (strain Spain)
C:Species: pepper mild mottle virus
C>Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 09-Jul-2004
C:Accession: JQ1315
R:Alonso, E.; Garcia-Luque, I.; de la Cruz, A.; Wicke, B.; Avila-Rincon, M.J.; Serra, M.
J. Gen. Virol. 72, 2875-2884, 1991
A/Title: Nucleotide sequence of the genomic RNA of pepper mild mottle virus, a resistant
A/Reference number: JQ1312; MUID:92113528; PMID:1765765
A/Accession: JQ1315
A/Molecule type: genomic RNA
A/Residues: 1-157 <ALO>
A/Cross-references: UNIPROT:P29096; UNIPARC:UPI0000000361; GB:M81413; NID:g333293; PIDN
A/Note: The authors translated the codon AAT for residue 26 as Ala, CAA for residue 46
C:Superfamily: tobacco mosaic virus coat protein
C/Keywords: acetylated amino end; coat protein
F:2/Modified site: acetylated amino end (Ala) (in mature form) #status predicted
Query Match 58.1%; Score 43; DB 1; Length 157;
Best Local Similarity 69.2%; Pred. No. 2.7;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 3 VENPMNRLVAETL 15
DB 97 VENPQNPTTAETL 109
||||:|||||
RESULT 3
A87266
sensor histidine kinase/response regulator [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: A87266
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Winn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A/Title: Complete Genome Sequence of Caulobacter crescentus.
A/Reference number: A87249; MUID:21173698; PMID:11259647
A/Accession: A87266
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-514 <STO>
A/Cross-references: UNIPROT:Q9ABT2; UNIPARC:UPI000000C6F46; GB:AE005673; NID:g13421249; I
C/Genetics:
A/Gene: CC0138
Query Match 55.4%; Score 41; DB 2; Length 514;
Best Local Similarity 57.1%; Pred. No. 24;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 2 AVENPMNRLVAETL 15
DB 385 AEDNPINALLAETL 398
|:|:|:|:|:|:|
RESULT 4

JC6516
signal recognition particle protein Sec65 - yeast (*Yarrowia lipolytica*)
N:Alternate names: SRP19
C:Species: *Yarrowia lipolytica*, *Candida lipolytica*
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C:Accession: JG6516; S44157
R:Sanchez, M.; Beckerich, J.M.; Gallardin, C.; Dominguez, A.
Gene 203, 75-84, 1997
A:Title: Isolation and cloning of the *Yarrowia lipolytica* SEC65 gene, a component of the
A:Reference number: JC6516; MUID:98085978; PMID:9426009
A:Accession: JC6516
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-310 <SAN>
A:Cross-references: UNIPROT:P41922; UNIPARC:UPI0000000313; EMBL:Z22570; NID:G473182; PID

Query Match 54.1%; Score 40; DB 2; Length 310;
Best Local Similarity 58.3%; Pred. No. 20;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 AVENPMNRLVAE 13
||||| : : ||
Db 146 AVENPLGQTIAE 157

RESULT 5
E69878
RNA-binding Sun protein homolog yloM - *Bacillus subtilis*
C:Species: *Bacillus subtilis*
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 05-Oct-2004
C:Accession: E69878
R:Kunst, F.; Ogawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funa, S.; Galizzi, A.; Galle
lech, J.; Harwood, C.R.; Hensaut, A.; Hilbert, H.; Holtsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon
A:Authors: Schleich, S.; Schroeder, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terptrata, P.; Tognoni, A.; Tosato, V.; Uchiyama
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: E69878
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-447 <KUN>
A:Cross-references: UNIPROT:P94464; UNIPARC:UPI00000603AF; GB:Z99112; GB:AL009126; NID:G
A:Experimental source: strain 168
C:Genetics:
C:Gene: yloM
C:Superfamily: ribosomal RNA small subunit methyltransferase B

Query Match 54.1%; Score 40; DB 1; Length 447;
Best Local Similarity 53.8%; Pred. No. 30;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 AVENPMNRLVAET 14
||| : : ||
Db 136 AIEDPVRLAET 148

RESULT 6
AH3098
rhizobium/RTX toxin [imported] - *Agrobacterium tumefaciens* (strain C58, Dupont)
C:Species: *Agrobacterium tumefaciens*
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: AH3098
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, J

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AH3098
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1944 <KUR>
A:Cross-references: UNIPROT:Q8U7N7; UNIPARC:UPI0000164922; GB:AE008689; PIDN:AAL45206.1;
A:Experimental source: strain C58 (Dupont)
C:Genetics:
C:Gene: rzca
A:Map position: linear chromosome

Query Match 54.1%; Score 40; DB 2; Length 1944;
Best Local Similarity 53.3%; Pred. No. 1.6e+02;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 FAVENPMNRLVAETL 15
||| : : ||
Db 255 FTVQSENKVVAAATL 269

RESULT 7
A96188
probable phosphoesterase (EC 3.1.1.-) yvnB [imported] - *Agrobacterium tumefaciens* (strain
C:Species: *Agrobacterium tumefaciens*
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C:Accession: A96188
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.,
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tum*
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: A96188
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1990 <KUR>
A:Cross-references: UNIPROT:Q8U7N7; UNIPARC:UPI000000D256D; GB:AE007870; PIDN:AAK89027.1;
C:Genetics:
C:Gene: AGR L 909
A:Map position: linear chromosome

Query Match 54.1%; Score 40; DB 2; Length 1990;
Best Local Similarity 53.3%; Pred. No. 1.6e+02;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 FAVENPMNRLVAETL 15
||| : : ||
Db 301 FTVQSENKVVAAATL 315

RESULT 8
A60023
coat protein - *Odontoglossum ringspot virus*
C:Species: *Odontoglossum ringspot virus*
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C:Accession: A60023
R:Dubs, M.C.; van Regenmortel, M.H.V.
Arch. Virol. 115, 239-249, 1990
A:Title: *Odontoglossum ringspot virus* coat protein: sequence and antigenic comparisons
A:Reference number: A60023; MUID:91083518; PMID:2260922
A:Accession: A60023
A:Status: preliminary
A:Molecule type: genomic RNA
A:Residues: 1-158 <DUB>
A:Cross-references: UNIPROT:O12854; UNIPARC:UPI00000178706
C:Superfamily: tobacco mosaic virus coat protein

Query Match 52.7%; Score 39; DB 2; Length 158;

Best Local Similarity 61.5%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 VENPMNRLVAETL 15
|||||
Db 99 VENPQNPTTETL 111

RESULT 9
S14468
coat protein - odontoglossum ringspot virus
C:Species: odontoglossum ringspot virus
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C:Accession: S14468
R:Isomura, Y.; Matumoto, Y.; Murayama, A.; Chatani, M.; Inouye, N.; Ikegami, M.
submitted to the EMBL Data Library, October 1990
A:Description: Nucleotide sequence of coat protein gene of odontoglossum ringspot virus.
A:Reference number: S14468
A:Accession: S14468
A>Status: preliminary
A:Molecule type: genomic RNA
A:Residues: 1-158 <ISO>
A:Cross-references: UNIPROT:P03578; UNIPARC:UPI00001710C2; EMBL:X55295; NID:g62084; PIDN
C:Superfamily: tobacco mosaic virus coat protein

Query Match 52.7%; Score 39; DB 2; Length 158;
Best Local Similarity 61.5%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 VENPMNRLVAETL 15
|||||
Db 99 VENPQNPTTETL 111

RESULT 10
JQ1266
coat protein - odontoglossum ringspot virus
C:Species: odontoglossum ringspot virus
C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004
C:Accession: JQ1266
R:Isomura, Y.; Matumoto, Y.; Murayama, A.; Chatani, M.; Inouye, N.; Ikegami, M.
J. Gen. Virol. 72, 2247-2249, 1991
A:Title: Molecular cloning, sequencing and expression in Escherichia coli of the odontog
A:Reference number: JQ1265; MUID:91374025; PMID:1895062
A:Accession: JQ1266
A:Molecule type: genomic RNA
A:Residues: 1-158 <ISO>
A:Cross-references: UNIPROT:P03578; UNIPARC:UPI0000178705
C:Genetics:
A:Introns: 304/3
C:Superfamily: tobacco mosaic virus coat protein

Query Match 52.7%; Score 39; DB 2; Length 158;
Best Local Similarity 61.5%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 VENPMNRLVAETL 15
|||||
Db 99 VENPQNPTTETL 111

RESULT 11
T02666
germin-like protein 16 - rice
C:Species: Oryza sativa (rice)
C>Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004
C:Accession: T02666
R:Yun, C.H.; Park, J.H.; Lee, J.H.; Eun, M.Y.
submitted to the EMBL Data Library, January 1998
A:Description: Nucleotide sequence of rice germin-like protein.
A:Reference number: Z14693
A:Accession: T02666
A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA
A:Residues: 1-221 <YUN>
A:Cross-references: UNIPROT:O49181; UNIPARC:UPI00000A3D8D; EMBL:AF042489; NID:g2801802;
A:Experimental source: strain Nipponbare
C:Genetics:
A:Gene: glp16
C:Superfamily: germin

Query Match 52.7%; Score 39; DB 2; Length 221;
Best Local Similarity 46.7%; Pred. No. 21;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 FAVENPMNRLVAETL 15
|:|:|:|:|:|:|
Db 132 FVMSNPDNRLLSKVL 146
|:|:|:|:|:|:|

RESULT 12
JC5590
cycloartenol synthase (EC 5.4.99.8) - garden pea
C:Species: Pisum sativum (garden pea)
C>Date: 04-Sep-1998 #sequence_revision 04-Sep-1998 #text_change 09-Jul-2004
C:Accession: JC5590
R:Morita, M.; Shibuya, M.; Lee, M.S.; Sankawa, U.; Ebizuka, Y.
Biol. Pharm. Bull. 20, 770-775, 1997
A:Title: Molecular cloning of pea cDNA encoding cycloartenol synthase and its functional
A:Reference number: JC5590; MUID:97399291; PMID:9255418
A:Accession: JC5590
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-756 <MOR>
A:Cross-references: UNIPROT:O23909; UNIPARC:UPI00000AA343; GB:D89619; NID:g2627180; PIDN
C:Superfamily: yeast lanosterol synthase
C:Keywords: intramolecular transferase; isomerase

Query Match 52.7%; Score 39; DB 2; Length 756;
Best Local Similarity 63.6%; Pred. No. 83;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 FAVENPMNRLV 11
|||:|:|:|:|:|
Db 67 FAKENPMNEVL 77
|||:|:|:|:|:|

RESULT 13
ISECTB
DNA topoisomerase (ATP-hydrolyzing) (EC 5.99.1.3) chain B - Escherichia coli (strain K-
N:Alternate names: DNA gyrase B chain; type II DNA topoisomerase B chain
C:Species: Escherichia coli
C>Date: 30-Sep-1988 #sequence_revision 31-Oct-1997 #text_change 05-Oct-2004
C:Accession: D65172; A26444; A26953; C22168; A38344
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: D65172
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-804 <BLAT>
A:Cross-references: UNIPROT:P06982; UNIPARC:UPI00001681B5; GB:AE000447; GB:U000096; NID:
A:Experimental source: strain K-12, substrain MG1655
R:Adachi, T.; Mizuuchi, M.; Robinson, E.A.; Appella, E.; O'Dea, M.H.; Gellert, M.; Mizu
Nucleic Acids Res. 15, 771-784, 1987
A:Title: DNA sequence of the E. coli gyrB gene: application of a new sequencing strateg
A:Reference number: A93674; MUID:87146392; PMID:3029692
A:Accession: A26444
A:Molecule type: DNA
A:Residues: 1-384, 'A' 386-804 <ADA>
A:Cross-references: UNIPARC:UPI000016552C; GB:X04341; GB:X00870; NID:g41643; PIDN:CAA27
R:Menzel, R.; Gellert, M.
J. Bacteriol. 169, 1272-1278, 1987
A:Title: Fusions of the Escherichia coli gyrA and gyrB control regions to the galactoki

A:Reference number: A26953; MUID:87137287; PMID:3029031
A:Accession: A26953

A:Molecule type: DNA

A:Residues: 1-23 <MEN>

A:Cross-references: UNIPARC:UPI000016F1EF; GB:M15548; NID:gl146307; PIDN:AAA23949.1; PID:
R:Adachi, T.; Mizuuchi, K.; Menzel, R.; Gellert, M.

Nucleic Acids Res. 12, 6389-6395, 1984

A:Title: DNA sequence and transcription of the region upstream of the E. coli gyrB gene.

A:Reference number: A22168; MUID:84297235; PMID:6089112

A:Accession: C22168

A:Molecule type: DNA

A:Residues: 1-50, 'R', 52-55, 'L', 57-106 <AD2>

A:Cross-references: UNIPARC:UPI00001730B3

R:Tamura, J.K.; Gellert, M.

J. Biol. Chem. 265, 21342-21349, 1990

A:Title: Characterization of the ATP binding site on Escherichia coli DNA gyrase. Affini

A:Reference number: A38344; MUID:91065955; PMID:2174443

A:Accession: A38344

A:Molecule type: protein

A:Residues: 93-126, 'X', 128-129 <TAM>

A:Cross-references: UNIPARC:UPI00001730B4

A:Note: Lys-103 and Lys-110 were shown to bind covalently to the ATP analog and DNA gyra

C:Comment: DNA gyrase, which catalyzes DNA supercoiling and relaxing, is made up of two

C:Genetics:

A:Gene: gyrB

A:Map position: 83 min

C:Superfamily: Type II topoisomerase, subunit B

C:Keywords: antibiotic resistance; ATP; DNA supercoiling; isomerase

F, 93-129/Region: ATP-binding

Query Match 52.7%; Score 39; DB 1; Length 804;

Best Local Similarity 64.3%; Pred. No. 89;

Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 AVENPMNRLVAETL 15

||| ||| ||| |||

Db 348 AVEQQMNEALLAYL 361

RESULT 14

DNA gyrase subunit B GyrB [imported] - Escherichia coli (strain O157:H7, substrain RIMD

C:Species: Escherichia coli

C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 05-Oct-2004

C:Accession: B91208

R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.

gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.

DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno

A:Reference number: A99629; MUID:21156231; PMID:11258796

A:Accession: B91208

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-804 <HAY>

A:Cross-references: UNIPROT:P06982; UNIPARC:UPI000016592C; GB:BA000007; PIDN:BA38057.1;

A:Experimental source: strain O157:H7, substrain RIMD 0509952

C:Genetics:

A:Gene: EC84634

C:Superfamily: Type II topoisomerase, subunit B

Query Match 52.7%; Score 39; DB 2; Length 804;

Best Local Similarity 64.3%; Pred. No. 89;

Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 AVENPMNRLVAETL 15

||| ||| ||| |||

Db 348 AVEQQMNEALLAYL 361

RESULT 15

D86054

hypothetical protein gyrB [imported] - Escherichia coli (strain O157:H7, substrain EDL93

C:Species: Escherichia coli

C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 05-Oct-2004
C:Accession: D86054

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew

iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouis, K.; Apodaca,

Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: D86054

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-804 <STO>

A:Cross-references: UNIPARC:UPI000016598B; GB:AE005174; NID:gl2518538; PIDN:AA658896.1;

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: gyrB

C:Superfamily: Type II topoisomerase, subunit B

Query Match 52.7%; Score 39; DB 2; Length 804;

Best Local Similarity 64.3%; Pred. No. 89;

Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 AVENPMNRLVAETL 15

||| ||| ||| |||

Db 348 AVEQQMNEALLAYL 361

RESULT 16

T02591

germin-like protein 3 - rice (fragment)

C:Species: Oryza sativa (rice)

C:Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 09-Jul-2004

C:Accession: T02591

R:Membre, N.; Bernier, F.

submitted to the EMBL Data Library, November 1997

A:Reference number: Z14634

A:Accession: T02591

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-181 <MEM>

A:Cross-references: UNIPROT:O48999; UNIPARC:UPI00000AB2D7; EMBL:AF032973; NID:G2655288;

C:Genetics:

A:Gene: GER3

C:Superfamily: germin

Query Match 51.4%; Score 38; DB 2; Length 181;

Best Local Similarity 46.7%; Pred. No. 25;

Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 FAVENPMNRLVAETL 15

||| ||| ||| |||

Db 92 FVTSNPDNRLLSKVL 106

RESULT 17

T40694

probable rna maturation protein - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C:Accession: T40694

R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Bothe, G.; Pohl, T.

submitted to the EMBL Data Library, February 1999

A:Reference number: Z21945

A:Accession: T40694

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-773 <WOO>

A:Cross-references: UNIPROT:O94689; UNIPARC:UPI000006BF80; EMBL:AL035536; PIDN:CAB36867;

A:Experimental source: strain 972h-; cosmid c83

C:Genetics:

A:Gene: SPDB:SPBC83.05

A:Map position: 2

C:Superfamily: Saccharomyces cerevisiae RNA12 protein

Query Match 51.4%; Score 38; DB 2; Length 773;
 Best Local Similarity 63.6%; Pred. No. 1.3e+02;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 VENPMNRLVAE 13
 : ||| : |||
 Db 480 LSNPMQVRVAE 490

RESULT 18
 F83292
 C;Species: Pseudomonas aeruginosa (strain
 C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
 C;Accession: F83292
 R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kae, A.; Larbig, K.; Lim,
 J.; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
 gen.
 A;Reference number: AB2950; MUID:20437337; PMID:10984043
 A;Accession: F83292
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-786 <STO>
 A;Cross-references: UNIPROT:Q91019; UNIPARC:UPI00000C582A; GB:AE004709; GB:AE004091; NID
 C;Genetics:
 A;Experimental source: strain PA01
 A;Gene: PA2824

Query Match 51.4%; Score 38; DB 2; Length 786;
 Best Local Similarity 58.3%; Pred. No. 1.3e+02;
 Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 4 ENPMNRLVAE 15
 : ||| : |||
 Db 670 DNPVNLVAKGL 681

RESULT 19
 H97798
 C;Species: Rickettsia conorii
 C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
 C;Accession: H97798
 R;Garcia, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; Ro
 science 293, 2093-2098, 2001
 A;Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
 A;Reference number: A97700; MUID:21442074; PMID:11557893
 A;Accession: H97798
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-50 <KUR>
 A;Cross-references: UNIPROT:Q92HH9; UNIPARC:UPI00000CBED9; GB:AE006914; PIDN:AAL03330.1;
 C;Genetics:
 A;Gene: RC0792

Query Match 50.0%; Score 37; DB 2; Length 50;
 Best Local Similarity 46.2%; Pred. No. 8.9;
 Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 PAVENPMNRLVAE 13
 : ||| : |||
 Db 28 FIVQEPNRLVCD 40

RESULT 20
 A48418
 C;Species: Rattus norvegicus (Norway rat)
 C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
 C;Accession: A48418; S37641
 R;Uberla, K.; Li, W.Q.; Qin, Z.H.; Richter, G.; Raabe, T.; Diamantstein, T.; Blankenstein

Cytokine 3, 72-81, 1991
 A;Title: The rat interleukin-5 gene: characterization and expression by retroviral gene
 A;Reference number: A48418; MUID:91355638; PMID:1653053
 A;Accession: A48418
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-132 <UBE>
 A;Cross-references: UNIPROT:Q08125; UNIPARC:UPI000012D4D3; EMBL:X54419; NID:G313254; PII
 A;Experimental source: cell line TRS-1
 A;Note: sequence extracted from NCBI backbone (NCBI:63651, NCBI:63652)
 C;Superfamily: interleukin-5
 C;Keywords: cytokine; glycoprotein; growth factor; homodimer; lymphokine; T-cell
 F;1-17/Domain: signal sequence #status predicted <SIG>
 F;18-132/Product: interleukin-5 #status predicted <MAT>
 F;45-74, 88/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F;61/Disulfide bonds: interchain (to 103) #status predicted
 F;103/Disulfide bonds: interchain (to 61) #status predicted

Query Match 50.0%; Score 37; DB 1; Length 132;
 Best Local Similarity 57.1%; Pred. No. 27;
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 AVENPMNRLVAE 15
 : ||| : |||
 Db 19 AMEIPMTVVKETL 32

RESULT 21
 ICMS5
 C;Species: Mus musculus (house mouse)
 C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004
 C;Accession: S00807; A38758; A24898; JS0077; PH0102; B39881
 R;Campbell, H.D.; Sanderson, C.J.; Wang, Y.; Hori, Y.; Martinson, M.E.; Tucker, W.Q.J.;
 Eur. J. Biochem. 174, 345-352, 1988
 A;Title: Isolation, structure and expression of cDNA and genomic clones for murine eosin
 n-5.
 A;Reference number: S00807; MUID:88254802; PMID:3133208
 A;Accession: S00807
 A;Molecule type: DNA
 A;Residues: 1-133 <CAM>
 A;Cross-references: UNIPROT:P04401; UNIPARC:UPI0000026C25; EMBL:X06271; NID:G53685; PIDN:
 A;Accession: A38758
 A;Molecule type: mRNA
 A;Residues: 1-133 <CAM2>
 A;Cross-references: UNIPARC:UPI0000026C25; EMBL:X06270; NID:G52687; PIDN:CAA29606.1; PII
 R;Kinashi, T.; Harada, N.; Severinson, E.; Tanabe, T.; Sideras, P.; Konishi, M.; Azuma,
 Nature 324, 70-73, 1986
 A;Title: Cloning of complementary DNA encoding T-cell replacing factor and identity with
 A;Reference number: A24898; MUID:87065032; PMID:3024009
 A;Accession: A24898
 A;Molecule type: mRNA
 A;Residues: 1-133 <KIN>
 A;Cross-references: UNIPARC:UPI0000026C25; EMBL:X04601; NID:G54898; PIDN:CAA28266.1; PII
 R;Mizuta, T.R.; Tanabe, T.; Nakakubo, H.; Noma, T.; Honjo, T.
 Growth Factors 1, 51-57, 1988
 A;Title: Molecular cloning and structure of the mouse interleukin-5 gene.
 A;Reference number: JS0077; MUID:90180853; PMID:3078564
 A;Accession: JS0077
 A;Molecule type: DNA
 A;Residues: 1-133 <MIZ>
 A;Cross-references: UNIPARC:UPI0000026C25
 R;Takahashi, T.; Yamaguchi, N.; Mita, S.; Yamaguchi, Y.; Suda, T.; Tomimaga, A.; Kikuchi,
 Mol. Immunol. 27, 911-920, 1990
 A;Title: Structural comparison of murine T-cell (B151K12)-derived T-cell-replacing factor
 A;Reference number: PH0102; MUID:91015093; PMID:2215480
 A;Accession: PH0102
 A;Molecule type: protein
 A;Residues: 21-45, X'47 <TAK>
 A;Cross-references: UNIPARC:UPI00000173682
 R;Yokota, T.; Coffman, R.L.; Hagiwara, H.; Rennick, D.M.; Takebe, Y.; Yokota, K.; Gemma;
 de Vries, J.; Lee, P.D.; Arai, N.; Arai, K.

Proc. Natl. Acad. Sci. U.S.A. 84, 7388-7392, 1987
A;Title: Isolation and characterization of lymphokine cDNA clones encoding mouse and human interleukin-5
A;Reference number: A39881; MUID:88041112; PMID:2823259
A;Accession: B39881
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-14, 'AA', 15-133 <YOK>
A;Cross-references: UNIPARC:UPI0000173683
C;Genetics:
A;Introns: 47/3; 58/3; 101/3
C;Superfamily: interleukin-5
C;Keywords: cytokine; glycoprotein; growth factor; homodimer; lymphokine; T-cell
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-133/Product: interleukin-5 #status predicted <MAT>
F;46/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;62/Disulfide bonds: interchain (to 104) #status predicted
F;75, 89/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;104/Disulfide bonds: interchain (to 62) #status predicted

Query Match 50.0%; Score 37; DB 1; Length 133;
Best Local Similarity 57.1%; Pred. No. 27;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 AVENPMNRLVAETL 15
| : | | | : | | |
Db 20 AMEIPMSTVVKETL 33

RESULT 22
S34856
coat protein mutant - tobacco mosaic virus
C;Species: tobacco mosaic virus, TMV
C;Date: 04-Jun-1999 #sequence_revision 04-Jun-1999 #text_change 09-Jul-2004
A;Accession: S34856; S34857; S34858; S34859
R;Kuhlmann, U.; Sarkar, S.; Rohde, W.
Nucleic Acids Res. 21, 3325, 1993
A;Title: Capsid protein gene sequences of four tobacco mosaic virus strains defective for replication
A;Reference number: S34856; MUID:93341948; PMID:8341610
A;Accession: S34856
A;Molecule type: mRNA
A;Residues: 1-159 <KUHI>
A;Cross-references: UNIPROT:P03570; UNIPROT:Q9IEM1; UNIPARC:UPI00001710E
A;Accession: S34857
A;Molecule type: mRNA
A;Residues: 1-103, 'T', 105-137, 'R', 139-159 <KUH2>
A;Cross-references: UNIPARC:UPI00001710AE; EMBL:X70883; NID:g313823; PIDN:CAA50228.1; PIDN:CAA50229.1; PIDN:CAA50230.1
A;Accession: S34858
A;Molecule type: mRNA
A;Residues: 1-78, 'S', 80-103, 'T', 105-137, 'R', 139-148, 'Y', 150-159 <KUH3>
A;Cross-references: UNIPARC:UPI00000F5E10; EMBL:X70884; NID:g313827; PIDN:CAA50229.1; PIDN:CAA50230.1
A;Accession: S34859
A;Molecule type: mRNA
A;Residues: 1-10, 'S', 12-13, 'F', 15-72 <KUH4>
A;Cross-references: UNIPARC:UPI00000F6406; EMBL:X70885; NID:g313825; PIDN:CAA50230.1; PIDN:CAA50231.1
C;Comment: These sequences are from mutagenized strains.

Query Match 50.0%; Score 37; DB 4; Length 159;
Best Local Similarity 61.5%; Pred. No. 33;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 VENQANPMTAETL 15
| : | | | : | | |
Db 97 VENQANPMTAETL 109

RESULT 23
D71206
hypothetical protein PH1918 - Pyrococcus horikoshii
C;Species: Pyrococcus horikoshii
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 05-Oct-2004
A;Accession: D71206
R;Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekiguchi, M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi, M.

DNA Res. 5, 55-76, 1998
A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic archaeon, Pyrococcus horikoshii
A;Reference number: A71000; MUID:98344137; PMID:9679194
A;Accession: D71206
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-162 <KAW>
A;Cross-references: UNIPROT:O59581; UNIPARC:UPI0000063178; GB:AP000007; NID:g3236134; PI
A;Experimental source: strain OT3
A;Note: this accession replaces an interim accession for a sequence replaced by GenBank
C;Genetics:
A;Gene: PH1918
C;Superfamily: hypothetical protein MTH699

Query Match 50.0%; Score 37; DB 1; Length 162;
Best Local Similarity 61.5%; Pred. No. 34;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 AVENPMNRLVAET 14
| : | | | : | | |
Db 111 ASENPFQVIVAET 123

RESULT 24
H75019
hypothetical protein PAB1234 - Pyrococcus abyssi (strain Orsay)
C;Species: Pyrococcus abyssi
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
A;Accession: H75019
R;anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A;Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome structure
A;Reference number: A75001
A;Accession: H75019
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-162 <KAW>
A;Cross-references: UNIPROT:Q9UY14; UNIPARC:UPI000006348B; GB:AJ248288; GB:AL096836; NID
A;Experimental source: strain Orsay
C;Genetics:
A;Gene: PAB1234
C;Superfamily: hypothetical protein MTH699

Query Match 50.0%; Score 37; DB 2; Length 162;
Best Local Similarity 61.5%; Pred. No. 34;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 AVENPMNRLVAET 14
| : | | | : | | |
Db 111 ASENPFQVIVAET 123

RESULT 25
E85565
probable tRNA synthetase ybeK [imported] - Escherichia coli (strain O157:H7, substrain F)
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
A;Accession: E85565
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: E85565
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-311 <STO>
A;Cross-references: UNIPROT:Q8XBL8; UNIPARC:UPI00000D0597; GB:AE005174; NID:g12513558; P
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: ybeK
C;Superfamily: yaaF protein

Query Match 50.0%; Score 37; DB 2; Length 311;
 Best Local Similarity 50.0%; Pred. No. 70;
 Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 AVENPMNRLVAETL 15
 : : : : :
 Db 207 AIGNPVSTIVAE LL 220

RESULT 26
 B90715
 probable tRNA synthetase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0559592)
 C:Species: Escherichia coli
 C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
 C:Accession: B90715
 R:Hayashi, T.; Makino, K.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; Gasaavara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.
 DNA Res. 8, 11-22, 2001
 A>Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genome analysis
 A:Reference number: A99629; MUID:21156231; PMID:11258796
 A:Accession: B90715
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-311 <HAV>
 A:Cross-references: UNIPROT:Q8XBL8; UNIPARC:UPI00000D0597; GB:BA000007; PIDN:BA034113.1
 A:Experimental source: strain O157:H7, substrain RIMD 0509952
 C:Genetics:
 C:Superfamily: yaaF protein

Query Match 50.0%; Score 37; DB 2; Length 311;
 Best Local Similarity 50.0%; Pred. No. 70;
 Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 AVENPMNRLVAETL 15
 : : : : :
 Db 207 AIGNPVSTIVAE LL 220

RESULT 27
 A64800
 purine nucleosidase-related protein ybeK - Escherichia coli (strain K-12)
 C:Species: Escherichia coli
 C>Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
 C:Accession: A64800
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Cohen, A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A>Title: The complete genome sequence of Escherichia coli K-12.
 A:Reference number: A64720; MUID:97426617; PMID:9278503
 A:Accession: A64800
 A>Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-311 <BLAT>
 A:Cross-references: UNIPROT:P41409; UNIPARC:UPI000013A301; GB:AE000169; GB:U00096; NID:9278503
 A:Experimental source: strain K-12, substrain MG1655
 C:Genetics:
 C:Superfamily: ybeK

Query Match 50.0%; Score 37; DB 2; Length 311;
 Best Local Similarity 50.0%; Pred. No. 70;
 Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 AVENPMNRLVAETL 15
 : : : : :
 Db 207 AIGNPVSTIVAE LL 220

RESULT 28
 B42519

A18R protein - vaccinia virus (strain Copenhagen)
 C:Species: vaccinia virus
 A>Note: host Homo sapiens (man)
 C>Date: 09-Nov-1990 #sequence_revision 09-Nov-1990 #text_change 09-Jul-2004
 C:Accession: B42519
 R:Johnson, G.P.
 submitted to GenBank, June 1990
 A:Reference number: A33172
 A:Accession: B42519
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-493 <JOH>
 A:Cross-references: UNIPROT:P20534; UNIPARC:UPI0000013804F

Query Match 50.0%; Score 37; DB 2; Length 493;
 Best Local Similarity 42.9%; Pred. No. 1.2e+02;
 Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 2 AVENPMNRLVAETL 15
 : : : : :
 Db 300 SVDEPRNQLIDTL 313

RESULT 29
 S23429
 protein kinase ERK3 (EC 2.7.1.1-) - human
 N:Alternate names: extracellular signal-regulated kinase 3; protein kinase, 63K
 C:Species: Homo sapiens (man)
 C>Date: 03-May-1994 #sequence_revision 20-Feb-1995 #text_change 05-Oct-2004
 C:Accession: S23429; S21580
 R:Gonzalez, F.A.; Raden, D.L.; Rigby, M.R.; Davis, R.J.
 PDBS Lett. 304, 170-178, 1992
 A>Title: Heterogeneous expression of four MAP kinase isoforms in human tissues.
 A:Reference number: S23426; MUID:92316223; PMID:1319925
 A:Accession: S23429
 A:Molecule type: mRNA
 A:Residues: 1-557 <GON>
 A:Cross-references: UNIPROT:P31152; UNIPARC:UPI000012F175; EMBL:X59727; NID:G23902; PIDN:G23902

C:Genetics:
 A:Gene: GDB:PRK4; MAPK4; ERK3
 A:Cross-references: GDB:135680; OMIM:176949
 A:Map position: 18q12-18q21
 C:Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
 F:18-281/Domain: protein kinase homology <KIN>
 F:26-34/Region: protein kinase ATP-binding motif

Query Match 50.0%; Score 37; DB 2; Length 557;
 Best Local Similarity 77.8%; Pred. No. 1.4e+02;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 NPMNRLVAE 13
 : : : : :
 Db 296 NPMRLTAE 304

RESULT 30
 T41681
 probable serine/threonine-protein kinase - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C:Accession: T41681
 R:Wedler, H.; Duesterhoeft, A.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.
 submitted to the EMBL Data Library, September 1999
 A:Reference number: Z21742
 A:Accession: T41681
 A>Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 1-650 <WED>
 A:Cross-references: UNIPROT:Q9U085; UNIPARC:UPI00000697F2; EMBL:AL117183; PIDN:CAB54861
 A:Experimental source: strain 972h-; clone pl p1811
 C:Genetics:
 A:Gene: SPDB:SPCP1E11.02
 A:Map position: 3

A:Introns: 42/1; 85/3; 235/1
 C:Superfamily: yeast hypothetical protein YNL020c; protein kinase homology
 Query Match 50.0%; Score 37; DB 2; Length 650;
 Best Local Similarity 53.8%; Pred. No. 1.6e+02;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 2 AVENPMNRLVAET 14
 :||| | :|
 Db 533 SVENPONNISAPT 545

Search completed: December 17, 2005, 07:33:37
 Job time : 42 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2005, 07:21:22 ; Search time 226 Seconds

(without alignments)

46.827 Million cell updates/sec

Title: US-10-787-382-20

Perfect score: 74

Sequence: 1 FAVENPMNRLVAETL 15

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : UniProt_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	74	100.0	118	2 QFTV10 CANFA	Q9cv10 canis famil
2	74	100.0	134	1 IL5 CANFA	Q95j76 canis famil
3	63	85.1	134	1 IL5 HORSE	O02699 equus caball
4	60	81.1	134	1 IL5 FELCA	O77515 felis silve
5	60	81.1	134	2 Q9MYM5 PIG	Q9mym5 sus scrofa
6	60	81.1	134	2 Q9TSD7 FELCA	Q9tsd7 felis silve
7	55	74.3	132	1 IL5 SHEEP	Q28586 ovis aries
8	55	74.3	134	1 IL5 BOVIN	P52173 bos taurus
9	47	63.5	696	2 Q9KYA5 STRCO	Q9kyas streptomyce
10	46	62.2	132	1 IL5 SIGHI	Q9sei9 sigmodon hi
11	46	62.2	457	2 Q846W1 STRCM	Q846w1 streptomyce
12	44	59.5	139	1 IL5 MACEU	Q9xt91 macropus eu
13	44	59.5	419	2 Q840Z9 9ENTR	O08402 citrobacter
14	43	58.1	156	1 COAT PMMV2	Q9wdg5 pepper mild
15	43	58.1	156	1 COAT PMMV1	Q84843 pepper mild
16	43	58.1	156	1 COAT PMMVJ	P69509 pepper mild
17	43	58.1	156	1 COAT PMMVS	P69510 pepper mild
18	43	58.1	156	1 COAT TOMKO	Q85482 tobacco mosa
19	43	58.1	157	2 Q6LER3 9VIRU	Q6ler3 tobacco mos
20	43	58.1	157	2 Q76M57 9VIRU	Q76m57 pepper mild
21	43	58.1	157	2 Q780A7 9VIRU	Q780a7 pepper mild
22	43	58.1	157	2 Q80P27 9VIRU	Q80p27 pepper mild
23	43	58.1	157	2 Q91E32 9VIRU	Q91e32 pepper mild
24	43	58.1	157	2 Q91U97 9VIRU	Q91u97 pepper mild
25	43	58.1	157	2 Q91U98 9VIRU	Q91u98 pepper mild
26	43	58.1	157	2 Q91U99 9VIRU	Q91u99 pepper mild
27	43	58.1	157	2 Q91UA0 9VIRU	Q91ua0 pepper mild
28	43	58.1	157	2 Q91UA1 9VIRU	Q91ua1 pepper mild
29	43	58.1	158	1 COAT PMMV0	Q9wdg3 pepper mild
30	43	58.1	158	2 Q834B3 9VIRU	Q834b3 tobacco mos
31	43	58.1	160	2 Q4VFY0 9VIRU	Q4vfy0 tropical so

32	43	58.1	257	2	Q4LQM7 9BURK	Q4lqm7 burkholderi
33	43	58.1	2679	2	Q6OPH9 CAEBR	Q6oph9 caenorhabdi
34	42	56.8	109	2	Q4YST9 PLABE	Q4yst9 plasmodium
35	42	56.8	222	1	GL19 ARATH	Q9fmb0 arabidopsis
36	42	56.8	1088	2	Q52G72 MAGGR	Q52gj2 magnaporthe
37	41	55.4	164	2	Q94SD6 CASSA	Q94sd6 castanea sa
38	41	55.4	307	2	Q9VJ65 DROME	Q9vjt65 drosophila
39	41	55.4	320	2	Q8INX3 DROME	Q8inx3 drosophila
40	41	55.4	514	2	Q9ABT2 CAUCR	Q9abt2 caulobacter
41	41	55.4	583	2	Q8ZMS8 SALTU	Q8zms8 salmoneilla
42	41	55.4	602	2	Q8F5P8 LEPIN	Q8f5p8 leptospira
43	41	55.4	602	2	Q72Q99 LEPIC	Q72q99 leptospira
44	41	55.4	866	2	Q623F4 CAEBR	Q623f4 caenorhabdi
45	40	54.1	90	2	Q7Y520 BPR69	Q7y520 bacterioph
46	40	54.1	252	2	Q6HMV5 BACHK	Q6hmv5 bacillus th
47	40	54.1	310	1	SEC65 YARLI	P41922 yarrowia li
48	40	54.1	315	2	Q9R9L5 RHIME	Q9r9l5 rhizobium m
49	40	54.1	315	2	Q9X9B3 RHIME	Q9x9b3 rhizobium m
50	40	54.1	345	2	Q8GRD4 SPPIO	Q8grd4 borrelia du
51	40	54.1	351	2	Q89FF6 BRAJA	Q89ff6 bradyrhizob
52	40	54.1	383	2	Q5SLZ6 CRYNE	Q5slz6 cryptococcu
53	40	54.1	383	2	Q5K8S2 CRYNE	Q5k8s2 cryptococcu
54	40	54.1	405	2	Q8D4F3 VIBVU	Q8d4f3 vibrio vuln
55	40	54.1	435	2	Q7MFY2 VIBVU	Q7mfy2 vibrio vuln
56	40	54.1	447	1	RSMB_BACSU	P94464 bacillus su
57	40	54.1	502	2	Q988B0 RHIL0	Q988b0 rhizobium l
58	40	54.1	759	2	Q4XZV6 PLACH	Q4xzv6 plasmodium
59	40	54.1	761	2	Q4YNR3 PLABE	Q4ynr3 plasmodium
60	40	54.1	771	2	Q7RNA9 PLAYO	Q7rna9 plasmodium
61	40	54.1	862	2	Q8IL36 PLAF7	Q8il36 plasmodium
62	40	54.1	1944	2	Q8U7N7 AGRTS	Q8u7n7 agrobacteri
63	40	54.1	1952	2	Q61XQ2 CAEBR	Q61xq2 caenorhabdi
64	40	54.1	1990	2	Q7CUW1 AGRT5	Q7cuw1 agrobacteri
65	40	54.1	4915	2	Q6CJB6 KULUA	Q6cjb6 kluyveromyc
66	39	52.7	147	2	Q7QSS7 GIALA	Q7qss7 giardia lam
67	39	52.7	147	2	Q8T4N8 GIALA	Q8t4n8 giardia lam
68	39	52.7	149	2	Q8VL28 ECOLI	Q8vl28 escherichia
69	39	52.7	157	1	COAT ORSVJ	P03578 odontogloss
70	39	52.7	157	1	COAT ORSVJ	Q84136 odontogloss
71	39	52.7	157	2	Q76U57 9VIRU	Q76u57 odontogloss
72	39	52.7	158	2	Q5K647 9VIRU	Q5k647 odontogloss
73	39	52.7	158	2	Q5K648 9VIRU	Q5k648 odontogloss
74	39	52.7	158	2	Q5K6A9 9VIRU	Q5k6a9 odontogloss
75	39	52.7	158	2	Q5K6B1 9VIRU	Q5k6b1 odontogloss
76	39	52.7	158	2	Q5K6B3 9VIRU	Q5k6b3 odontogloss
77	39	52.7	158	2	Q5K6B6 9VIRU	Q5k6b6 odontogloss
78	39	52.7	158	2	Q5K6C1 9VIRU	Q5k6c1 odontogloss
79	39	52.7	158	2	Q5TUK4 9VIRU	Q5tjk4 odontogloss
80	39	52.7	158	2	Q5TUK5 9VIRU	Q5tjk5 odontogloss
81	39	52.7	158	2	Q80P28 9VIRU	Q80p28 odontogloss
82	39	52.7	158	2	Q80P29 9VIRU	Q80p29 odontogloss
83	39	52.7	158	2	Q80P30 9VIRU	Q80p30 odontogloss
84	39	52.7	158	2	Q12854 9VIRU	Q12854 odontogloss
85	39	52.7	172	2	Q6W159 RHISN	Q6w159 rhizobium s
86	39	52.7	180	2	Q9U249 CAEEL	Q9u249 caenorhabdi
87	39	52.7	221	2	Q49181 ORYSA	Q49181 oryza sativ
88	39	52.7	252	2	Q81HD1 BACCR	Q81hd1 bacillus ce
89	39	52.7	278	2	Q6MAX1 PARCW	Q6max1 parachlamyd
90	39	52.7	323	2	Q937H5 HAFAL	Q937h5 hafnia alve
91	39	52.7	364	2	Q4UPL9 XANCP	Q4upl9 xanthomonas
92	39	52.7	364	2	Q8P437 XANCP	Q8p437 xanthomonas
93	39	52.7	369	2	Q754V8 ASHGO	Q754v8 ashbya goss
94	39	52.7	390	2	Q8LOR2 ECOLI	Q8lor2 escherichia
95	39	52.7	390	2	Q8LOR3 ECOLI	Q8lor3 escherichia
96	39	52.7	390	2	Q8LOR4 ECOLI	Q8lor4 escherichia
97	39	52.7	390	2	Q8KHV8 ECOLI	Q8khv8 escherichia
98	39	52.7	390	2	Q8KIA4 ECOLI	Q8kia4 escherichia
99	39	52.7	390	2	Q8KIA0 SHISO	Q8kia0 shigella so
100	39	52.7	390	2	Q8KIA1 SHIFL	Q8kia1 shigella fl
101	39	52.7	390	2	Q8KIA2 SHIBO	Q8kia2 shigella bo
102	39	52.7	390	2	Q8KIA3 ECOLI	Q8kia3 escherichia
103	39	52.7	390	2	Q70GL9 RETFI	Q70gl9 reticulomyc
104	39	52.7	408	2		

105 39 52.7 502 2 Q5YUM9 NOCFA
 106 39 52.7 518 2 Q9VF72 DROME
 107 39 52.7 576 2 Q5CN36 CRYPO
 108 39 52.7 586 2 Q5CYA3 CRYPO
 109 39 52.7 692 2 Q9AD7 RHIZO
 110 39 52.7 756 2 Q23909 PEA
 111 39 52.7 757 2 Q9SKV6 GLYGL
 112 39 52.7 803 1 GYEB EGOLI
 113 39 52.7 805 2 Q8FV7 ECOL6
 114 39 52.7 829 2 Q4PD86 USTLAGO
 115 39 52.7 832 2 Q7YYP6 CRYPO
 116 39 52.7 844 2 Q5CWS4 CRYPO
 117 39 52.7 949 2 Q5B369 EMENI
 118 39 52.7 1104 2 Q75F16 ASHGO
 119 38.5 52.0 234 2 Q80LPS NPVAH
 120 38.5 52.0 717 2 Q87J33 NEUPORA
 121 38 51.4 85 2 Q93EH8 HELHP
 122 38 51.4 135 2 Q6NKC1 CORDI
 123 38 51.4 181 2 Q48999 ORYSA
 124 38 51.4 221 2 Q6YZA8 ORYSA
 125 38 51.4 223 2 Q6GQH8 XENLA
 126 38 51.4 235 2 Q4MXH7 RHOSH
 127 38 51.4 252 2 Q4MWH7 BACCE
 128 38 51.4 252 2 Q73CV7 BACCI
 129 38 51.4 252 2 Q81UK1 BACAN
 130 38 51.4 274 2 Q62D39 BURMA
 131 38 51.4 274 2 Q63M77 BURPS
 132 38 51.4 344 2 Q75DGI ASHGO
 133 38 51.4 401 2 Q82NG0 STRAW
 134 38 51.4 471 1 TYPH MOUSE
 135 38 51.4 521 2 Q7VJ33 HELHP
 136 38 51.4 528 2 Q8G9V8 ECOLI
 137 38 51.4 538 2 Q5P5Q6 AZOSE
 138 38 51.4 617 2 Q6M987 NEUCR
 139 38 51.4 623 2 Q6S7E9 MOUSE
 140 38 51.4 631 2 Q65M44 BACLD
 141 38 51.4 714 2 Q7SA10 NEUCR
 142 38 51.4 721 2 Q511J9 ENTHI
 143 38 51.4 728 2 Q6MYU3 ASPFU
 144 38 51.4 766 2 Q4WSR2 ASPFU
 145 38 51.4 773 2 Q94689 SCHPO
 146 38 51.4 786 2 Q91019 PSEAE
 147 38 51.4 923 2 Q4J1A7 AZOVI
 148 38 51.4 937 2 Q822R7 ENTHI
 149 38 51.4 939 2 Q4WXJ6 ASPFU
 150 38 51.4 1041 2 Q8L816 CATRO

ALIGNMENTS

RESULT 1
 Q9TV10 CANFA PRELIMINARY; PRT; 118 AA.
 AC Q9TV10
 DT 01-MAY-2000 (T-EMBLrel. 13, Created)
 DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
 DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
 DE Interleukin-5 (Fragment).
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
 OC Canis.
 NCBI_TaxID=9615;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA German A.J., Helps C.R., Harley R., Hall E.J., Day M.J.;
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Factor that induces terminal differentiation of late-
 CC developing B-cells to immunoglobulin secreting cells (By
 CC similarity).
 CC -!- SUBUNIT: Homodimer; disulfide-linked (By similarity).
 CC EMBL; AF091133; AAD46991.1; -; mRNA.

DR HSSP; P05113; 1HUL.
 DR Ensembl; ENSCAFG0000000855; Canis familiaris.
 DR GO; GO:0005576; C:extracellular region; IEA.
 DR GO; GO:0005137; F:interleukin-5 receptor binding; IEA.
 DR GO; GO:0006955; P:immune response; IEA.
 DR InterPro; IPR000186; Interleukin_5.
 DR Pfam; PF02025; IL5; 1.
 DR PRINTS; PR00432; INTERLEUKIN5.
 DR ProDom; PD006721; Interleukin_5; 1.
 FT NON_TER 1 118
 SQ SEQUENCE 118 AA; 13507 MW; 36A5563DD67C968C CRC64;
 Query Match 100.0%; Score 74; DB 2; Length 118;
 Best Local Similarity 100.0%; Pred.No. 5.4e-05; Indels 0; Gaps 0;
 Matches 15; Conservative 0; Mismatches 0;
 QY 1 FAVENPMNRLVAETL 15
 DB 8 FAVENPMNRLVAETL 22
 ID IL5 CANFA STANDARD; PRT; 134 AA.
 AC Q95J76;
 DT 29-MAR-2004 (Rel. 43, Created)
 DT 29-MAR-2004 (Rel. 43, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Interleukin-5 precursor (IL-5) (T-cell replacing factor) (TRF)
 DE (Eosinophil differentiation factor).
 GN Name=IL5;
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
 OC Canis.
 NCBI_TaxID=9615;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=21334408; PubMed=11440633; DOI=10.1089/107999001750277835;
 RA Yang S., Sellins K.S., Weber E., McCall C.;
 RT "Canine interleukin-5: molecular characterization of the gene and
 RT expression of biologically active recombinant protein.";
 RL J. Interferon Cytokine Res. 21:361-367(2001).
 CC -!- FUNCTION: Factor that induces terminal differentiation of late-
 CC developing B-cells to immunoglobulin secreting cells (By
 CC similarity).
 CC -!- SUBUNIT: Homodimer; disulfide-linked (By similarity).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the IL-5 family.
 CC
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC
 CC EMBL; AF331920; AAL10716.1; -; Genomic_DNA.
 DR EMBL; AF331919; AAL10715.1; -; mRNA.
 DR HSSP; P05113; 1HUL.
 DR Ensembl; ENSCAFG0000000855; Canis familiaris.
 DR InterPro; IPR012351; Cytokine_4_hlx.
 DR InterPro; IPR000186; Interleukin_5.
 DR PANTHER; PTHR10525; Interleukin_5; 1.
 DR Pfam; PF02025; IL5; 1.
 DR PRINTS; PR00432; INTERLEUKIN5.
 DR ProDom; PD006721; Interleukin_5; 1.
 KW Cytokine; Glycoprotein; Growth factor; Signal.
 FT SIGNAL 1 21 By similarity.
 FT CHAIN 22 134 Interleukin-5.
 FT CARBOHYD 76 76 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 90 90 N-linked (GlcNAc...) (Potential).
 FT DISULFID 63 63 Interchain (with C-105) (By similarity).

```

FT DISULFID 105 105 Interchain (with C-63) (By similarity).
SQ SEQUENCE 134 AA; 15307 MW; 003C86D94D6FF4C6 CRC64;

Query Match 100.0%; Score 74; DB 1; Length 134;
Best Local Similarity 100.0%; Pred. No. 6.2e-05; Mismatches 0; Indels 0; Gaps 0;
Matches 15; Conservative 0;

QY 1 FAVENPMNRLVAETL 15
Db 20 FAVENPMNRLVAETL 34

RESULT 3
IL5_HORSE
ID IL5_HORSE STANDARD; PRT; 134 AA.
AC 002699.
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Interleukin-5 precursor (IL-5) (T-cell replacing factor) (TRF)
DE (Eosinophil differentiation factor).
GN Name=IL5;
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RA Vandersgrift E.V., Horobov D.W.;
RT "Equine interleukin-5.";
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Factor that induces terminal differentiation of late-
CC developing B-cells to immunoglobulin secreting cells (By
CC similarity).
CC -!- SUBUNIT: Homodimer; disulfide-linked (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the IL-5 family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; U91947; AAB51382.1; -; mRNA.
CC SMR; 002699; 25-130.
CC InterPro; IPR012351; Cytokine 4 hlx.
CC InterPro; IPR000186; Interleukin_5.
CC PANTHER; PTHR10525; Interleukin_5; 1.
CC Pfam; PF02025; IL5; 1.
CC PRINTS; PR00432; INTERLEUKINS.
CC ProDom; PD006721; Interleukin_5; 1.
CC Cytokine; Glycoprotein; Growth factor; Signal.
CC SIGNAL 1 19
CC CHAIN 20 134
CC CARBOHYD 76 76 N-linked (GlcNAc...) (Potential).
CC CARBOHYD 90 90 N-linked (GlcNAc...) (Potential).
CC DISULFID 63 63 Interchain (with C-105) (By similarity).
CC DISULFID 105 105 Interchain (with C-63) (By similarity).
CC PRINTS; PR00432; INTERLEUKINS.
CC ProDom; PD006721; Interleukin_5; 1.
CC Cytokine; Glycoprotein; Growth factor; Signal.
CC SIGNAL 1 19
CC CHAIN 20 134
CC CARBOHYD 76 76 N-linked (GlcNAc...) (Potential).
CC CARBOHYD 90 90 N-linked (GlcNAc...) (Potential).
CC DISULFID 63 63 Interchain (with C-105) (By similarity).
CC DISULFID 105 105 Interchain (with C-63) (By similarity).
CC SEQUENCE 134 AA; 15081 MW; 2914840E3E7A006D CRC64;

Query Match 85.1%; Score 63; DB 1; Length 134;
Best Local Similarity 92.9%; Pred. No. 0.0057; Mismatches 1; Indels 0; Gaps 0;
Matches 13; Conservative 1;

QY 2 AVENPMNRLVAETL 15
Db 21 AVESPMNRLVAETL 34

RESULT 4
IL5_FELCA

```

```

ID IL5_FELCA STANDARD; PRT; 134 AA.
AC 077515; 062740;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Interleukin-5 precursor (IL-5) (T-cell replacing factor) (TRF)
DE (Eosinophil differentiation factor).
GN Name=IL5;
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
OC Felinae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98452719; PubMed=9781459;
RA Padrid P.A., Qin Y., Wells T.N.C., Solway J., Camoretti-Mercado B.;
RT "Sequence and structural analysis of feline interleukin-5 cDNA.";
RL Am. J. Vet. Res. 59:1263-1269(1998).
RN [2]
RP NUCLEOTIDE SEQUENCE OF 12-128.
RA Harley R., Day M.J., Gruffydd-Jones T.J., Harbour D.A., Helps C.R.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Factor that induces terminal differentiation of late-
CC developing B-cells to immunoglobulin secreting cells (By
CC similarity).
CC -!- SUBUNIT: Homodimer; disulfide-linked (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the IL-5 family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; AF025436; AAC64505.1; -; mRNA.
CC EMBL; AF051372; AAC05752.1; -; mRNA.
CC HSSP; P05113; 1HUL.
CC InterPro; IPR012351; Cytokine 4 hlx.
CC InterPro; IPR000186; Interleukin_5.
CC PANTHER; PTHR10525; Interleukin_5; 1.
CC Pfam; PF02025; IL5; 1.
CC PRINTS; PR00432; INTERLEUKINS.
CC ProDom; PD006721; Interleukin_5; 1.
CC Cytokine; Glycoprotein; Growth factor; Signal.
CC SIGNAL 1 19
CC CHAIN 20 134
CC CARBOHYD 76 76 N-linked (GlcNAc...) (Potential).
CC CARBOHYD 90 90 N-linked (GlcNAc...) (Potential).
CC DISULFID 63 63 Interchain (with C-105) (By similarity).
CC DISULFID 105 105 Interchain (with C-63) (By similarity).
CC CONFLICT 104 105 KC -> NF (in Ref. 2).
CC CONFLICT 108 111 ERWR -> KKWK (in Ref. 2).
CC CONFLICT 114 114 K -> N (in Ref. 2).
CC CONFLICT 117 117 D -> N (in Ref. 2).
CC CONFLICT 121 121 V -> F (in Ref. 2).
CC CONFLICT 125 126 VI -> LL (in Ref. 2).
CC SEQUENCE 134 AA; 15224 MW; 87D18B8F8C820 CRC64;

Query Match 81.1%; Score 60; DB 1; Length 134;
Best Local Similarity 85.7%; Pred. No. 0.02; Mismatches 2; Indels 0; Gaps 0;
Matches 12; Conservative 2;

QY 2 AVENPMNRLVAETL 15
Db 21 AVQSPMNRLVAETL 34

RESULT 5
Q9MYM5_PIG
ID Q9MYM5_PIG PRELIMINARY; PRT; 134 AA.
AC Q9MYM5;

```

DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
 DE Interleukin-5.
 GB Name=IL-5;
 GB Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
 OC Sus.
 NCBI_TaxID=9823;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=20130134; PubMed=10663563; DOI=10.1007/s002510050009;
 RA Sylvain H., Matvienko O., Leonchiks A., Aiving K., van der Ploeg I.;
 RT "Molecular cloning, expression, and purification of pig interleukin-5";
 RL Immunogenetics 51:59-64(2000).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC Tissue=Lung;
 RA Johnsen C.K., Grondahl-Hansen J., Johansen R., Jørgensen G.,
 RA Heegaard P.M.H.;
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Factor that induces terminal differentiation of late-developing B-cells to immunoglobulin secreting cells (By similarity).
 CC -!- SUBUNIT: Homodimer; disulfide-linked (By similarity).
 CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
 DR EMBL; AJ010088; CAB70611.2; -; mRNA.
 DR EMBL; AJ133452; CAB38328.1; -; mRNA.
 DR HSP; P05113; 1HUL.
 DR SMR; Q9NVM5; 29-130.
 DR GO; GO:0005576; C:extracellular region; IEA.
 DR GO; GO:0005125; F:Cytokine activity; IEA.
 DR GO; GO:0005137; F:interleukin-5 receptor binding; IEA.
 DR GO; GO:0006955; P:immune response; IEA.
 DR InterPro; IPR000186; Interleukin_5.
 DR Pfam; PF02025; IL5; 1.
 DR PRINTS; PD00432; INTERLEUKINS.
 DR ProDom; PD006721; Interleukin_5; 1.
 KW Cytokine; Growth factor.
 SQ SEQUENCE 134 AA; 15191 MW; B485D562A028A899 CRC64;
 BR
 Query Match 81.1%; Score 60; DB 2; Length 134;
 Best Local Similarity 92.9%; Pred.No. 0.02; Mismatches 0; Indels 0; Gaps 0;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 AVENPMNRLVAETL 15
 DB 21 AVENTWNLVAETL 34
 BR
 RESULT 6
 Q9TSD7 FELCA PRELIMINARY; PRT; 134 AA.
 ID Q9TSD7
 AC Q9TSD7
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Interleukin 5.
 OS Felis silvestris catus (Cat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
 OC Felis.
 NCBI_TaxID=9685;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Vandegriff E., Hughes K.J., O'Reilly K.L.;
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Factor that induces terminal differentiation of late-developing B-cells to immunoglobulin secreting cells (By similarity).
 CC -!- SUBUNIT: Homodimer; disulfide-linked (By similarity).

CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
 DR EMBL; AF068770; AAC27616.1; -; mRNA.
 DR HSP; P05113; 1HUL.
 DR GO; GO:0005576; C:extracellular region; IEA.
 DR GO; GO:0005125; F:Cytokine activity; IEA.
 DR GO; GO:0005137; F:interleukin-5 receptor binding; IEA.
 DR GO; GO:0006955; P:immune response; IEA.
 DR InterPro; IPR000186; Interleukin_5.
 DR Pfam; PF02025; IL5; 1.
 DR PRINTS; PD00432; INTERLEUKINS.
 DR ProDom; PD006721; Interleukin_5; 1.
 KW Cytokine; Growth factor.
 SQ SEQUENCE 134 AA; 15176 MW; 9A118B78F8C820 CRC64;
 BR
 Query Match 81.1%; Score 60; DB 2; Length 134;
 Best Local Similarity 85.7%; Pred.No. 0.02; Mismatches 2; Indels 0; Gaps 0;
 Matches 12; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 2 AVENPMNRLVAETL 15
 DB 21 AVQSPMNLVAETL 34
 BR
 RESULT 7
 IL5 SHEEP STANDARD; PRT; 132 AA.
 ID IL5 SHEEP
 AC Q28586;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Interleukin-5 precursor (IL-5) (T-cell replacing factor) (TRF)
 DE (Eosinophil differentiation factor).
 GN Name=IL5;
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
 OC Pecora; Bovidae; Caprinae; Ovis.
 OX NCBI_TaxID=9940;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Bryson C.E., Viney E., Brandon M., Boyd A.W.;
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Seow H.-P., David M.-J., McWaters P.G., Hurst L., Wood P.R.;
 RT "Cloning of ovine interleukin-5 cDNA";
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Factor that induces terminal differentiation of late-developing B-cells to immunoglobulin secreting cells (By similarity).
 CC -!- SUBUNIT: Homodimer; disulfide-linked (By similarity).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the IL-5 family.
 CC
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.
 CC
 DR EMBL; U17053; AAB60629.1; -; Genomic DNA.
 DR EMBL; U17052; AAB60629.1; JOINED; Genomic DNA.
 DR EMBL; U35038; AAC59991.1; -; mRNA.
 DR HSP; P05113; 1HUL.
 DR InterPro; IPR012351; Cytokine_4_hlx.
 DR InterPro; IPR000186; Interleukin_5.
 DR PANTHER; PTHR10525; Interleukin_5; 1.
 DR Pfam; PF02025; IL5; 1.
 DR PRINTS; PR00432; INTERLEUKINS.
 DR ProDom; PD006721; Interleukin_5; 1.
 KW Cytokine; Glycoprotein; Growth factor; Signal.
 FT SIGNAL 1 19 By similarity.
 FT CHAIN 20 132 Interleukin-5.

FT CARBOHYD 74 74 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 88 88 N-linked (GlcNAc...) (Potential).
 FT DISULFID 61 61 Interchain (with C-103) (By similarity).
 FT DISULFID 103 103 Interchain (with C-61) (By similarity).
 SQ SEQUENCE 132 AA; 14974 MW; D783F2B720E249D9 CRC64;

Query Match 74.3%; Score 55; DB 1; Length 132;
 Best Local Similarity 85.7%; Pred. No. 0.15;
 Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 AVENPMNRLVAETL 15
 DB 19 AVESTMNRVAETL 32

RESULT 8

IL5_BOVIN STANDARD; PRT; 134 AA.
 AC P52173;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Interleukin-5 precursor (IL-5) (T-cell replacing factor) (TRF)
 DE (Eosinophil differentiation factor).
 GN Name=IL5;
 OS Bos taurus (Bovina).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
 OC Pecora; Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Blood;
 RX MEDLINE=97075944; PubMed=8918267; DOI=10.1016/0378-1119(96)00252-1;
 RA Mertens B., Gibrig E., Seow H.F.;
 RT "The nucleotide sequence of the bovine interleukin-5-encoding cDNA."
 RL Gene 176:273-274(1996).
 CC -!- FUNCTION: Factor that induces terminal differentiation of late-developing B-cells to immunoglobulin secreting cells (By similarity).
 CC -!- SUBUNIT: Homodimer; disulfide-linked.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the IL-5 family.

 This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

 DR EMBL; Z67872; CAA91779.1; -; mRNA.
 DR PIR; JC5116; JC5116.
 DR HSP; P05113; 1HUL.
 DR InterPro; IPR012351; Cytokine_4_hlx.
 DR InterPro; IPR000186; Interleukin_5.
 DR PANTHER; PTHR10525; Interleukin_5.
 DR Pfam; PF02025; IL5; 1.
 DR PRINTS; PR00432; INTERLEUKIN5.
 DR ProDom; PD006721; Interleukin_5; 1.
 KW Cytokine; Glycoprotein; Growth factor; Signal.
 FT SIGNAL 1 21 By similarity.
 FT CHAIN 22 134 Interleukin-5.
 FT CARBOHYD 76 76 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 90 90 N-linked (GlcNAc...) (Potential).
 FT DISULFID 63 63 Interchain (with C-105) (By similarity).
 FT DISULFID 105 105 Interchain (with C-63) (By similarity).
 SQ SEQUENCE 134 AA; 15229 MW; 0B3A2328EE7431F4 CRC64;

Query Match 74.3%; Score 55; DB 1; Length 134;
 Best Local Similarity 85.7%; Pred. No. 0.15;
 Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 AVENPMNRLVAETL 15

DB 21 AVESTMNRVAETL 34

RESULT 9

Q9KYA5_STRCO PRELIMINARY; PRT; 696 AA.
 ID Q9KYA5;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Putative O-methyltransferase.
 GN OrderedLocNames=SCO6928; ORFNames=SC1B2.34c;
 OS Streptomyces coelicolor.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.

RC STRAIN=A3(2) / M145;
 RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
 RA Bentley S.D., Chater K.F., Cerdono-Tarraga A.-M., Challis G.L., Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H., Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M., Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S., Huang C.-H., Kieser T., Larke L., Murphy L.D., Oliver K., O'Neill S., Rabinowitsch E., Rajandream M.A., Rutherford K.M., Rutter S., Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K., Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J., Hopwood D.A.;
 RA "Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2).";
 RT Nature 417:141-147(2002).
 RL EMBL; AL939129; CAB92590.1; -; Genomic_DNA.
 DR HSP; Q872R3; 1JGL.
 DR GO; GO:0004719; P:protein-L-isopartate (D-aspartate) O-meth. .; IEA.
 DR GO; GO:0016740; P:transferase activity; IEA.
 DR GO; GO:0006464; P:protein modification; IEA.
 DR InterPro; IPR000682; PCMT.
 DR InterPro; IPR000051; SAM_bind.
 DR Pfam; PF01135; PCMT; 1.
 KW Complete proteome; Methyltransferase; Transferase.
 SQ SEQUENCE 696 AA; 75886 MW; 5C08C043308CFCBA CRC64;

Query Match 63.5%; Score 47; DB 2; Length 696;

Best Local Similarity 61.5%; Pred. No. 23;
 Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 FAVENPMNRLVAR 13
 DB 682 FAFDNLNRIIVIE 694

RESULT 10

IL5_SIGHI STANDARD; PRT; 132 AA.
 AC Q985I9;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Interleukin-5 precursor (IL-5) (T-cell replacing factor) (TRF)
 DE (Eosinophil differentiation factor).
 GN Name=IL5;
 OS Sigmodon hispidus (Hispid cotton rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Cricetidae; Sigmodontinae; Sigmodon.
 OX NCBI_TaxID=42415;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Spleen;

RX PubMed=11054577; DOI=10.1016/S0378-1119(00)00366-8;

QY Houard S., Jacquet A., Haumont M., Damiet V., Milican F., Glineur F.,

```

RA Bollen A.;
RT "Cloning, expression and purification of recombinant cotton rat
RL interleukin-5.";
RT Gene 257:149-155(2000).
CC -!- FUNCTION: Factor that induces terminal differentiation of late-
CC developing B-cells to immunoglobulin secreting cells.
CC -!- SUBUNIT: Homodimer; disulfide-linked.
CC -!- SUBCELLULAR LOCATION: secreted.
CC -!- SIMILARITY: Belongs to the IL-5 family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
DR EMBL; AF148211; AAG16722.1; -; mRNA.
DR HSSP; P05113; 1HUL.
DR SMR; Q9BSI9; 22-128.
DR InterPro; IPR012351; Cytokine 4 hlx.
DR InterPro; IPR000186; Interleukin_5.
DR PANTHER; PTHR10525; Interleukin_5; 1.
DR Pfam; PF02025; IL5; 1.
DR PRINTS; PR00432; INTERLEUKIN5.
DR ProDom; PD006721; Interleukin_5; 1.
KW Cytokine; Glycoprotein; Growth factor; Signal.
FT SIGNAL 1 19 By similarity.
FT CHAIN 20 132 Interleukin-5.
FT CARBOHYD 45 45 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 74 74 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 88 88 N-linked (GlcNAc... ) (Potential).
FT DISULFID 61 61 Interchain (with C-103) (By similarity).
FT DISULFID 103 103 Interchain (with C-61) (By similarity).
SQ SEQUENCE 132 AA; 15434 MW; B328B81B2371FEB9 CRC64;

Query Match 62.2%; Score 46; DB 1; Length 132;
Best Local Similarity 66.7%; Pred. No. 6.1;
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 FAVENPMNRLVAETL 15
DB 18 FAVEIPMHTVVKETL 32
|||||:|:|:|

RESULT 11
Q846W1_STRCM
ID Q846W1_STRCM PRELIMINARY; PRT; 457 AA.
AC Q846W1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
OS Streptomyces cinnamomeus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1900;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC15413;
RX MEDLINE=22823264; PubMed=12940979;
RX DOI=10.1046/j.1365-2958.2003.03571.x;
RA Olinyk M., Stark C.B.W., Bhatt A., Jones M.A., Hughes-Thomas Z.A.,
RA Wilkinson C., Olinyk Z., Demychuk Y., Staunton J., Leadlay P.F.;
RA "Analysis of the biosynthetic gene cluster for the polyether
RT antibiotic monensin in Streptomyces cinnamomeus and evidence for the
RT role of monB and monC genes in oxidative cyclization.";
RL Mol. Microbiol. 49:1179-1190(2003).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC15413;
RA Olinyk M.;
RT "The gene cluster for monensin biosynthesis.";
RL Thesis (1999), University of Cambridge.

```

```

DR EMBL; AF440781; AA065811.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 457 AA; 50456 MW; 9D53DB3362009F54 CRC64;

Query Match 62.2%; Score 46; DB 2; Length 457;
Best Local Similarity 69.2%; Pred. No. 22;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 FAVENPMNRLVAE 13
DB 191 FSVHEPMSRLVAE 203
|:|:|:|:|:|:|

RESULT 12
IL5_MACEU
ID IL5_MACEU STANDARD; PRT; 139 AA.
AC Q9XT91;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Interleukin-5 precursor (IL-5) (T-cell replacing factor) (TRF)
DE (Eosinophil differentiation factor).
DE Name=IL5;
OS Macropus eugenii (Tamar wallaby).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.
OX NCBI_TaxID=9315;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=99432005; PubMed=10501836; DOI=10.1007/s002510050577;
RA Hawken R.J., Maccarone P., Toder R., Marshall Graves J.A.,
RA Maddox J.P.;
RT "Isolation and characterization of marsupial IL5 genes.";
RL Immunogenetics 49:942-948(1999).
CC -!- FUNCTION: Factor that induces terminal differentiation of late-
CC developing B-cells to immunoglobulin secreting cells (By
CC similarity).
CC -!- SUBUNIT: Homodimer; disulfide-linked (By similarity).
CC -!- SUBCELLULAR LOCATION: secreted.
CC -!- SIMILARITY: Belongs to the IL-5 family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
DR EMBL; AF064209; AAD37462.1; -; Genomic_DNA.
DR HSSP; P05113; 1HUL.
DR InterPro; IPR012351; Cytokine 4 hlx.
DR InterPro; IPR000186; Interleukin_5.
DR PANTHER; PTHR10525; Interleukin_5; 1.
DR Pfam; PF02025; IL5; 1.
DR PRINTS; PR00432; INTERLEUKIN5.
DR ProDom; PD006721; Interleukin_5; 1.
KW Cytokine; Glycoprotein; Growth factor; Signal.
FT SIGNAL 1 19 By similarity.
FT CHAIN 20 139 Interleukin-5.
FT CARBOHYD 48 48 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 77 77 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 91 91 N-linked (GlcNAc... ) (Potential).
FT DISULFID 64 64 Interchain (with C-106) (By similarity).
FT DISULFID 106 106 Interchain (with C-64) (By similarity).
SQ SEQUENCE 139 AA; 15784 MW; CE16342A68F10622 CRC64;

Query Match 59.5%; Score 44; DB 1; Length 139;
Best Local Similarity 64.3%; Pred. No. 15;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 AVENPMNRLVAETL 15
DB 22 ATGNPVSRLVTETL 35
|:|:|:|:|:|:|

```



```

RESULT 13
O08402_9ENTR PRELIMINARY; PRT; 419 AA.
ID O08402_9ENTR PRELIMINARY; PRT; 419 AA.
AC O08402_009479;
DT 01-JUL-1997 (TRENBLrel. 04, Created)
DT 01-JUL-1997 (TRENBLrel. 04, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE DNA gyrase beta-subunit (Fragment).
GN Name=gyrB;
OS Citrobacter sp. JYME-1.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Citrobacter.
OX NCBI_TaxID=60487;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=JYME-1;
RA Venkateswaran K., Nealeon K.H.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBSJ databases.
CC -!- SIMILARITY: Belongs to the type II topoisomerase family.
DR EMBL; AF005700; AAB0843.1; -; Genomic_DNA.
DR HSSP; P06982; IE11.
DR SMR; O08402; 1-286.
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0003918; P:DNA topoisomerase (ATP-hydrolyzing) activity; IEA.
DR GO; GO:0006265; P:DNA topological change; IEA.
DR InterPro; IPR011558; DNA gyrase_B.
DR InterPro; IPR001241; DNA topoisomII.
DR InterPro; IPR000565; DNA topoisomIV_B.
DR Pfam; PF00204; DNA gyraseB; 1.
DR PRINTS; PR00418; TP12FAMILY.
DR ProDom; PD149633; DNA_gyrase_B; 1.
DR SMART; SM00433; TOP2c; 1.
DR PROSITE; PS00177; TOPOISOMERASE II; 1.
KW ATP-binding; Isomerase; Nucleotide-binding; Topoisomerase.
FT NON_TER 1 419
FT NON_TER 419 419
SQ SEQUENCE 419 AA; 46232 MW; E442532D303FDA52 CRC64;

Query Match 59.5%; Score 44; DB 2; Length 419;
Best Local Similarity 64.3%; Pred. No. 46;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 AVENPMNRLVAETL 15
||| ||| :|||
Db 257 AVEQPMNLLSEYL 270

RESULT 14
COAT PMMV2
ID COAT PMMV2 STANDARD; PRT; 156 AA.
AC Q9WDG5;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Coat protein.
GN Name=CP;
OS Pepper mild mottle virus (strain P2) (PMMV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
OX NCBI_TaxID=138305;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RA Sohn S.-H., Hahn J.-H., Hwang Y.-S.;
RT "The coat protein gene of Pepper Mild Mottle Virus isolated from Hot pepper in Korea."
RL Submitted (NOV-1998) to the EMBL/GenBank/DBSJ databases.
CC -!- SIMILARITY: Belongs to the tobamoviruses coat protein family.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

Query Match 58.1%; Score 43; DB 1; Length 156;
Best Local Similarity 69.2%; Pred. No. 25;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 3 VENPMNRLVAETL 15
||| | |||
Db 96 VENQNPPTAETL 108

RESULT 16
COAT PMMVJ
ID COAT PMMVJ STANDARD; PRT; 156 AA.
AC P69509; P29096;
DT 01-DEC-1992 (Rel. 24, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)

```

```

CC EMBL; AF103778; AAD20289.1; -; mRNA.
DR HSSP; P03579; 1VTM.
DR InterPro; IPR001337; TMV coat.
DR Pfam; PF00721; TMV coat; 1.
KW Acetylation; Capsid protein; Structural protein.
FT INIT_MET 0
FT MOD_RES 1 0 N-acetylalanine (By similarity).
SQ SEQUENCE 156 AA; 17125 MW; ECD6A3AD176A919 CRC64;

Query Match 58.1%; Score 43; DB 1; Length 156;
Best Local Similarity 69.2%; Pred. No. 25;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 3 VENPMNRLVAETL 15
||| | |||
Db 96 VENQNPPTAETL 108

RESULT 15
COAT PMMVI
ID COAT PMMVI STANDARD; PRT; 156 AA.
AC Q84843;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Coat protein.
GN Name=CP;
OS Pepper mild mottle virus (strain Italian) (PMMV-I).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
OX NCBI_TaxID=138306;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC RNA].
RX MEDLINE=93319451; PubMed=8328918;
RA Garcia-Luque I., Ferrero M.L., Rodriguez J.M., Alonso E.,
de la Cruz A., Sanz A.I., Vaquero C., Serra M.T., Diaz-Ruiz J.R.;
RT "The nucleotide sequence of the coat protein genes and 3' non-coding regions of two resistance-breaking tobamoviruses in pepper shows that they are different viruses."
RL Arch. Virol. 131:75-88(1993).
CC -!- SIMILARITY: Belongs to the tobamoviruses coat protein family.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

EMBL; X72587; CRA51184.1; -; Genomic_RNA.
DR HSSP; P03570; 2TMV.
DR SMR; Q84843; 1-156.
DR InterPro; IPR001337; TMV coat.
DR Pfam; PF00721; TMV coat; 1.
KW Acetylation; Capsid protein; Structural protein.
FT INIT_MET 0
FT MOD_RES 1 0 N-acetylalanine (By similarity).
SQ SEQUENCE 156 AA; 17135 MW; F9F50AFD5952CF06 CRC64;

Query Match 58.1%; Score 43; DB 1; Length 156;
Best Local Similarity 69.2%; Pred. No. 25;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 3 VENPMNRLVAETL 15
||| | |||
Db 96 VENQNPPTAETL 108

RESULT 16
COAT PMMVJ
ID COAT PMMVJ STANDARD; PRT; 156 AA.
AC P69509; P29096;
DT 01-DEC-1992 (Rel. 24, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)

```

```

DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Coat protein.
GN Name=CP;
OS Pepper mild mottle virus (strain Japan) (PMMV-J).
VIRUSES; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
OC NCBI_TaxID=138663;
RN [1]_
RP NUCLEOTIDE SEQUENCE [MRNA].
RA Kirita M., Akutsu K., Watanabe Y., Tsuda S.;
RT "Nucleotide sequence of the Japanese isolate of pepper [Capsicum
annuum] mild mottle tobamovirus (TMV-P) RNA.";
RL Ann. Phytopathol. Soc. Jpn. 63:373-376(1997).
CC -!- SIMILARITY: Belongs to the tobamoviruses coat protein family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
CC
DR EMBL; AB000709; BAA19169.1; -; mRNA.
DR HSSP; P03570; 2TMV.
DR SMR; P69509; 1-156.
DR InterPro; IPR001337; TMV coat.
DR Pfam; PF00721; TMV coat; 1.
KW Acetylation; Capsid protein; Structural protein.
FT INIT_MET 0 By similarity.
FT MOD_RES 1 1 N-acetylalanine (By similarity).
SQ SEQUENCE 156 AA; 17110 MW; B8D2E3E7C955BF73 CRC64;

Query Match 58.1%; Score 43; DB 1; Length 156;
Best Local Similarity 69.2%; Pred. No. 25;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 VENPMNRLVAETL 15
Db 96 VENPQNPTTAETL 108

RESULT 17
COAT_PMMVS
ID COAT_PMMVS STANDARD; PRT; 156 AA.
AC P69570; P29096;
DT 01-DEC-1992 (Rel. 24, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Coat protein.
GN Name=CP;
OS Pepper mild mottle virus (strain Spain) (PMMV-S).
VIRUSES; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
OC NCBI_TaxID=31745;
RN [1]_
RP NUCLEOTIDE SEQUENCE [GENOMIC RNA].
RX MEDLINE=92113528; PubMed=1765765;
RA Alonso E., Garcia-Luque I., de la Cruz A., Wicke B.;
RA Avila-Rincon M.J., Serra M.T., Castresana C., Diaz-Ruiz J.R.;
RT "Nucleotide sequence of the genomic RNA of pepper mild mottle virus, a
resistance-breaking tobamovirus in pepper.";
RL J. Gen. Virol. 72:2875-2884(1991).
CC -!- SIMILARITY: Belongs to the tobamoviruses coat protein family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
CC
DR EMBL; M81413; AAB02337.1; -; Genomic_RNA.
DR FIR; JQ3315; VCTMPV.
DR HSSP; P03570; 2TMV.
DR SMR; P69510; 1-156.
DR InterPro; IPR001337; TMV coat.
DR Pfam; PF00721; TMV_coat; 1.

```

```

KW Acetylation; Capsid protein; Structural protein.
FT INIT_MET 0 By similarity.
FT MOD_RES 1 1 N-acetylalanine (By similarity).
SQ SEQUENCE 156 AA; 17110 MW; B8D2E3E7C955BF73 CRC64;

Query Match 58.1%; Score 43; DB 1; Length 156;
Best Local Similarity 69.2%; Pred. No. 25;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 VENPMNRLVAETL 15
Db 96 VENPQNPTTAETL 108

RESULT 18
COAT_TOMKO
ID COAT_TOMKO STANDARD; PRT; 156 AA.
AC Q83482;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Coat protein.
GN Name=CP;
OS Tomato mosaic virus (strain Korean) (ToMV).
VIRUSES; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
OC NCBI_TaxID=138313;
RN [1]_
RP NUCLEOTIDE SEQUENCE [GENOMIC RNA].
RA Park E.K., Lee C.H., Lee Y.G., Lee Y.H.;
RT "Characterization of coat protein from TMV Korean tomato strain.";
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the tobamoviruses coat protein family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
CC
DR EMBL; J35074; AAA46589.1; -; Genomic_RNA.
DR HSSP; P03570; 2TMV.
DR SMR; Q83482; 1-156.
DR InterPro; IPR001337; TMV coat.
DR Pfam; PF00721; TMV coat; 1.
KW Acetylation; Capsid protein; Structural protein.
FT INIT_MET 0 By similarity.
FT MOD_RES 1 1 N-acetylserine (By similarity).
SQ SEQUENCE 156 AA; 17126 MW; B8CB04736FEAF41 CRC64;

Query Match 58.1%; Score 43; DB 1; Length 156;
Best Local Similarity 69.2%; Pred. No. 25;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 VENPMNRLVAETL 15
Db 96 VENPQNPTTAETL 108

RESULT 19
Q6LER3_9VIRU
ID Q6LER3_9VIRU PRELIMINARY; PRT; 157 AA.
AC Q6LER3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Coat protein.
OS Tobacco mosaic virus.
VIRUSES; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
OC NCBI_TaxID=12242;
RN [1]_
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=P;
RA Nishimiya S.;

```

RL Submitted (OCT-1992) to the EMBL/GenBank/DBJ databases.
 DR EMBL; D13367; BAA02631.1; -; Genomic_RNA.
 DR SMR; Q6LER3; 2-157.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR001337; TMV_coat.
 DR Pfam; PF00721; TMV_coat; 1.
 KW Capsid protein.
 SQ SEQUENCE 157 AA; 17241 MW; 7122837E9E86257 CRC64;

Query Match 58.1%; Score 43; DB 2; Length 157;
 Best Local Similarity 69.2%; Pred. No. 25;
 Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 VENPMNRLVAETL 15
 |||||
 Db 97 VENPQNPTTAETL 109

RESULT 20

ID Q76M57_9VIRU PRELIMINARY; PRT; 157 AA.
 AC Q76M57
 DT 05-JUL-2004 (TRENBLrel. 27, Created)
 DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
 DT 01-FEB-2005 (TRENBLrel. 29, Last annotation update)
 DE Coat protein (Capsid protein).
 OS Pepper mild mottle virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
 OX NCBI_TaxID=12239;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Hamada H., Takeuchi S., Kiba A., Hikichi Y., Okuno T.;
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22255451; PubMed=12367724; DOI=10.1016/S0166-0934(02)00135-0;
 RA Letschert B., Adam G., Lesemann D.E., Willingmann P., Heinze C.;
 RT "Detection and differentiation of serologically cross-reacting
 RT tobamoviruses of economical importance by RT-PCR and RT-PCR-RFLP.";
 RL J. Virol. Methods 106:1-10(2002).
 DR EMBL; AB062053; BAB5800.1; -; Genomic_RNA.
 DR EMBL; AJ429089; CAD22086.1; -; Genomic_RNA.
 DR SMR; Q76M57; 2-157.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR001337; TMV_coat.
 DR Pfam; PF00721; TMV_coat; 1.
 KW Capsid protein.
 SQ SEQUENCE 157 AA; 17266 MW; 30056A6479EF1222 CRC64;

Query Match 58.1%; Score 43; DB 2; Length 157;
 Best Local Similarity 69.2%; Pred. No. 25;
 Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 VENPMNRLVAETL 15
 |||||
 Db 97 VENPQNPTTAETL 109

RESULT 21

ID Q780A7_9VIRU PRELIMINARY; PRT; 157 AA.
 AC Q780A7
 DT 05-JUL-2004 (TRENBLrel. 27, Created)
 DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
 DT 10-MAY-2005 (TRENBLrel. 30, Last annotation update)
 DE Coat protein.
 GN Name=CP;
 OS Pepper mild mottle virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
 OX NCBI_TaxID=12239;
 RN [1]

RP NUCLEOTIDE SEQUENCE.
 RA Yoon J.Y., Choi J.K., Ryu K.H.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Takeuchi S., Hamada H., Kiba A., Hikichi Y., Okuno T.;
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RA Eiras M.;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP NUCLEOTIDE SEQUENCE.
 RA Huang Y.Y., Zhai X.L., Ma R.Q.;
 RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=Kr;
 RA Yoon J., Choi J., Ryu K.;
 RT "Molecular characterization of pepper mild mottle virus Kr strain.";
 RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=PMOV-CN;
 RA Liu F., Wang X., Zhou G.;
 RT "Nucleotide sequence analysis of the genome of the pepper mild mottle
 RT virus in China (PMOV).";
 RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB084456; BAB91239.1; -; Genomic_RNA.
 DR EMBL; AB119482; BAC85000.1; -; Genomic_RNA.
 DR EMBL; AF525080; AAM83090.1; -; mRNA.
 DR EMBL; AB126003; BAD90601.1; -; Genomic_RNA.
 DR EMBL; AY859497; AAW55641.1; -; Genomic_RNA.
 DR EMBL; AY632863; AAT46356.1; -; Genomic_RNA.
 DR SMR; Q780A7; 2-157.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR001337; TMV_coat.
 DR Pfam; PF00721; TMV_coat; 1.
 KW Capsid protein.
 SQ SEQUENCE 157 AA; 17241 MW; 7122837E9E86257 CRC64;

Query Match 58.1%; Score 43; DB 2; Length 157;
 Best Local Similarity 69.2%; Pred. No. 25;
 Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 VENPMNRLVAETL 15
 |||||
 Db 97 VENPQNPTTAETL 109

RESULT 22

ID Q80P27_9VIRU PRELIMINARY; PRT; 157 AA.
 AC Q80P27;
 DT 01-JUN-2003 (TRENBLrel. 24, Created)
 DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Capsid protein.
 OS Pepper mild mottle virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
 OX NCBI_TaxID=12239;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22255451; PubMed=12367724; DOI=10.1016/S0166-0934(02)00135-0;
 RA Letschert B., Adam G., Lesemann D.E., Willingmann P., Heinze C.;
 RT "Detection and differentiation of serologically cross-reacting
 RT tobamoviruses of economical importance by RT-PCR and RT-PCR-RFLP.";
 RL J. Virol. Methods 106:1-10(2002).
 DR EMBL; AJ429087; CAD22084.1; -; Genomic_RNA.
 DR HSSP; P03570; 2TMV.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.

```
DR InterPro; IPR001337; TMV_coat.
DR Pfam; PF00721; TMV_coat; 1.
SQ SEQUENCE 157 AA; 17310 MW; 47703D64B776081A CRC64;

Query Match 58.1%; Score 43; DB 2; Length 157;
Best Local Similarity 69.2%; Pred. No. 25;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 VENPMNRLVAETL 15
Db 97 VENPQNPTTAETL 109

RESULT 23
Q91E32_9VIRU
ID Q91E32_9VIRU PRELIMINARY; PRT; 157 AA.
AC Q91E32;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Coat protein.
OS Pepper mild mottle virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
OX NCBI_TaxID=12239;
RN [1]_
RP STRAIN:IA;
RC MEDLINE=22255465; PubMed=12367738; DOI=10.1016/S0166-0934(02)00144-1;
RX Velasco L., Janssen D., Ruiz-Garcia L., Segundo E., Cuadrado I.M.;
RA "The complete nucleotide sequence and development of a differential
RT detection assay for a pepper mild mottle virus (PMoV) isolate that
RT overcomes L3 resistance in pepper.";
RL J. Virol. Methods 106:135-140(2002).
DR EMBL; AJ308226; CACS9958.1; -; Genomic_RNA.
DR HSSP; P03570; 2TMV.
DR SMR; Q91E32; 2-157.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001337; TMV_coat.
DR Pfam; PF00721; TMV_coat; 1.
KW Capsid protein.
SQ SEQUENCE 157 AA; 17252 MW; D2C9E9DC426CB9AA CRC64;

Query Match 58.1%; Score 43; DB 2; Length 157;
Best Local Similarity 69.2%; Pred. No. 25;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 VENPMNRLVAETL 15
Db 97 VENPQNPTTAETL 109

RESULT 24
Q91U97_9VIRU
ID Q91U97_9VIRU PRELIMINARY; PRT; 157 AA.
AC Q91U97;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Coat protein (Capsid protein).
OS Pepper mild mottle virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
OX NCBI_TaxID=12239;
RN [1]_
RP NUCLEOTIDE SEQUENCE.
RA Hamada H., Takeuchi S., Kiba A., Hikichi Y., Okuno T.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RN [2]_
RP NUCLEOTIDE SEQUENCE.
RA Hamada H., Takeuchi S., Kiba A., Hikichi Y., Okuno T.;
RX MEDLINE=22255451; PubMed=12367724; DOI=10.1016/S0166-0934(02)00135-0;
RT Letscher B., Adam G., Lesemann D.E., Willingmann P., Heinze C.;
RT "Detection and differentiation of serologically cross-reacting
RT tobamoviruses of economical importance by RT-PCR and RT-PCR-RELP.";

J. Virol. Methods 106:1-10(2002).
DR EMBL; AB062054; BAB5801.1; -; Genomic_RNA.
DR EMBL; AJ429088; CAD22085.1; -; Genomic_RNA.
DR HSSP; P03570; 2TMV.
DR SMR; Q91U97; 2-157.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001337; TMV_coat.
DR Pfam; PF00721; TMV_coat; 1.
KW Capsid protein.
SQ SEQUENCE 157 AA; 17241 MW; 7472D5980A5ED257 CRC64;

Query Match 58.1%; Score 43; DB 2; Length 157;
Best Local Similarity 69.2%; Pred. No. 25;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 VENPMNRLVAETL 15
Db 97 VENPQNPTTAETL 109

RESULT 25
Q91U98_9VIRU
ID Q91U98_9VIRU PRELIMINARY; PRT; 157 AA.
AC Q91U98;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Coat protein.
OS Pepper mild mottle virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
OX NCBI_TaxID=12239;
RN [1]_
RP NUCLEOTIDE SEQUENCE.
RA Hamada H., Takeuchi S., Kiba A., Hikichi Y., Okuno T.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB062052; BAB55799.1; -; Genomic_RNA.
DR HSSP; P03570; 2TMV.
DR SMR; Q91U98; 2-157.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001337; TMV_coat.
DR Pfam; PF00721; TMV_coat; 1.
KW Capsid protein.
SQ SEQUENCE 157 AA; 17255 MW; BBAP2818317760CA CRC64;

Query Match 58.1%; Score 43; DB 2; Length 157;
Best Local Similarity 69.2%; Pred. No. 25;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 VENPMNRLVAETL 15
Db 97 VENPQNPTTAETL 109

RESULT 26
Q91U99_9VIRU
ID Q91U99_9VIRU PRELIMINARY; PRT; 157 AA.
AC Q91U99;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Coat protein.
OS Pepper mild mottle virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
OX NCBI_TaxID=12239;
RN [1]_
RP NUCLEOTIDE SEQUENCE.
RA Hamada H., Takeuchi S., Kiba A., Hikichi Y., Okuno T.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB062051; BAB55798.1; -; Genomic_RNA.
DR HSSP; P03570; 2TMV.
DR SMR; Q91U99; 2-157.
```

DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR001337; TMV_coat.
 DR Pfam; PF00721; TMV_coat; 1.
 KW Capsid protein.
 SQ SEQUENCE 157 AA; 17236 MW; E90E1098121A6BBB CRC64;

Query Match 58.1%; Score 43; DB 2; Length 157;
 Best Local Similarity 69.2%; Pred. No. 25;
 Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 VENPMNRLVAETL 15
 |||||
 Db 97 VENPQNPTTAETL 109

RESULT 27

Q91UA0_9VIRU Q91UA0_9VIRU PRELIMINARY; PRT; 157 AA.
 ID Q91UA0; AC Q91UA0;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Coat protein.
 OS Pepper mild mottle virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
 OX NCBI_TaxID=12239;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Hanada H., Takeuchi S., Kiba A., Hikichi Y., Okuno T.;
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB062050; BAB55797.1; -; Genomic_RNA.
 DR HSSP; P03570; 2TMV.
 DR SMR; Q91UA0; 2-157.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR001337; TMV_coat.
 DR Pfam; PF00721; TMV_coat; 1.
 KW Capsid protein.
 SQ SEQUENCE 157 AA; 17226 MW; A93C0CFB2F493B6B CRC64;

Query Match 58.1%; Score 43; DB 2; Length 157;
 Best Local Similarity 69.2%; Pred. No. 25;
 Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 VENPMNRLVAETL 15
 |||||
 Db 97 VENPQNPTTAETL 109

RESULT 28

Q91UA1_9VIRU Q91UA1_9VIRU PRELIMINARY; PRT; 157 AA.
 ID Q91UA1; AC Q91UA1;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Coat protein.
 OS Pepper mild mottle virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
 OX NCBI_TaxID=12239;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Hanada H., Takeuchi S., Kiba A., Hikichi Y., Okuno T.;
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB062049; BAB55796.1; -; Genomic_RNA.
 DR HSSP; P03570; 2TMV.
 DR SMR; Q91UA1; 2-157.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR001337; TMV_coat.
 DR Pfam; PF00721; TMV_coat; 1.
 KW Capsid protein.

SQ SEQUENCE 157 AA; 17210 MW; A80E0CE80E761BA7 CRC64;

Query Match 58.1%; Score 43; DB 2; Length 157;
 Best Local Similarity 69.2%; Pred. No. 25;
 Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 VENPMNRLVAETL 15
 |||||
 Db 97 VENPQNPTTAETL 109

RESULT 29

COAT_PMV0 COAT_PMV0 STANDARD; PRT; 158 AA.
 ID COAT_PMV0; AC Q9WDG3;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Coat protein.
 GN Name=CP;
 OS Pepper mild mottle virus (strain P0) (PMV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
 OX NCBI_TaxID=138303;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [MRNA].
 RA Sohn S.-H., Hahn J.-H., Hwang Y.-S.;
 RT "The coat protein gene of Pepper Mild Mottle Virus isolated from Hot pepper in Korea";
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the tobamoviruses coat protein family.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.
 CC EMBL; AF103776; AAD20287.1; -; mRNA.
 DR HSSP; P03570; 2TMV.
 DR SMR; Q9WDG3; 1-158.
 DR InterPro; IPR001337; TMV_coat.
 DR Pfam; PF00721; TMV_coat; 1.
 KW Acetylation; Capsid protein; Structural protein.
 FT INTMET 0 By similarity.
 FT MOD_RES 1 1 N-acetylalanine (By similarity).
 SQ SEQUENCE 158 AA; 17451 MW; A31C592C0A01D79A CRC64;

Query Match 58.1%; Score 43; DB 1; Length 158;
 Best Local Similarity 69.2%; Pred. No. 25;
 Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 VENPMNRLVAETL 15
 |||||
 Db 96 VENPQNPTTAETL 108

RESULT 30

Q83483_9VIRU Q83483_9VIRU PRELIMINARY; PRT; 158 AA.
 ID Q83483; AC Q83483;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Coat protein.
 OS Tobacco mosaic virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
 OX NCBI_TaxID=12242;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Park E.K., Lee C.H., Lee Y.G., Lee Y.H.;
 RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
 DR EMBL; L35073; AAA46588.1; -; Genomic_RNA.
 DR HSSP; P03570; 2TMV.

DR SMR; Q83483; 2-148.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; P:structural molecule activity; IEA.
DR InterPro; IPR001337; TMV_coat.
DR Pfam; PF00721; TMV_coat; 1.
KW Capsid protein.
SQ SEQUENCE 158 AA; 17456 MW; 8F4A2E610E5F41A5 CRC64;

Query Match 58.1%; Score 43; DB 2; Length 158;
Best Local Similarity 69.2%; Pred. No. 25;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 VENPMNRLVAETL 15
||| |
Db 97 VENPQNPTTAETL 109

Search completed: December 17, 2005, 07:32:57
Job time : 233 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 21, 2005, 16:35:41 ; Search time 3047 Seconds
(without alignments)
279.833 Million cell updates/sec

Title: US-10-787-382-20

Perfect score: 74
Sequence: 1 FAVENPMRLVAETL 15

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO.epool/US10787382/runat_16122005_132436_8012/app query fasta_1.199
-DB=GenEmbl -QWIT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=500
-DOCALIGN=200 -THR SCORE=spt -THR MAX=100 -THR MIN=0 -ALIGN=50 -MODE=LOCAL
-OUTPMT=spt -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10787382 @CEN 1.1 4939 @runat_16122005_132436_8012 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb.ba:*
2: gb.in:*
3: gb.env:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pr:*
9: gb.ro:*
10: gb.sts:*
11: gb.sy:*
12: gb.un:*
13: gb.vi:*
14: gb.htg:*
15: gb.pli:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	74	100.0	345	6	BD211562 Canine an
2	74	100.0	345	6	BD211563 Canine an
3	74	100.0	345	6	AR241540 Sequence

c	4	74	100.0	345	6	AR241541	Sequence
c	5	74	100.0	345	6	AR254496	Sequence
c	6	74	100.0	345	6	AR254497	Sequence
c	7	74	100.0	356	4	AF091133	Canis fam
c	8	74	100.0	402	6	BD211560	Canine an
c	9	74	100.0	402	6	BD211561	Canine an
c	10	74	100.0	402	6	AR241538	Sequence
c	11	74	100.0	402	6	AR241539	Sequence
c	12	74	100.0	402	6	AR254494	Sequence
c	13	74	100.0	402	6	AR254495	Sequence
c	14	74	100.0	405	6	AR300436	Sequence
c	15	74	100.0	405	6	AX083939	Sequence
c	16	74	100.0	610	4	AF331919	Canis fam
c	17	74	100.0	610	6	BD211558	Canine an
c	18	74	100.0	610	6	BD211559	Canine an
c	19	74	100.0	610	6	AR241536	Sequence
c	20	74	100.0	610	6	AR241537	Sequence
c	21	74	100.0	610	6	AR254492	Sequence
c	22	74	100.0	610	6	AR254493	Sequence
c	23	74	100.0	1658	4	AF331920	Canis fam
c	24	63	85.1	405	4	ECU91947	Equus caball
c	25	63	85.1	172134	14	AC158733	Rhinoloph
c	26	63	85.1	177136	14	AC158734	Rhinoloph
c	27	60	81.1	354	4	AF051372	Felis cat
c	28	60	81.1	405	4	AF068770	Felis cat
c	29	60	81.1	405	4	SSC010088	Sus scrof
c	30	60	81.1	529	4	SSC133452	Sus scrof
c	31	60	81.1	838	4	AF025436	Felis cat
c	32	55	74.3	405	4	BTINTLEUS	B. taurus mR
c	33	55	74.3	520	4	OU35038	Ovis aries
c	34	55	74.3	1140	4	OALV1	Ovis aries
c	35	55	74.3	197131	4	AC149665	Bos taurus
c	36	49	66.2	110000	1	CR543861	Continuation (15 o
c	37	47	63.5	1247	15	AF419552	Arabidops
c	38	47	63.5	106329	15	AC004411	Arabidops
c	39	46	62.2	242	13	AF336001	Streptomy
c	40	46	62.2	448	9	AF148211	Streptomy
c	41	46	62.2	13600	6	AX250264	Sequence
c	42	46	62.2	16535	14	AC014950	Sequence
c	43	46	62.2	17525	2	AB071139	Apis mell
c	44	46	62.2	103450	1	AF440781	Streptomy
c	45	46	62.2	110000	1	CP000086	Continuation (31 o
c	46	46	62.2	186241	2	AC007086	Drosophila
c	47	46	62.2	201020	8	AC022809	Homo sapi
c	48	46	62.2	226575	14	AC164180	Bos tauru
c	49	46	62.2	249287	14	AC161834	Bos tauru
c	50	46	62.2	254954	14	AC156428	Bos tauru
c	51	46	62.2	297107	2	AB003833	Drosophila
c	52	46	60.8	997	1	AF314230	Mycoplasma
c	53	45	60.8	110000	15	CR382130	Continuation (30 o
c	54	45	60.8	134151	15	AC006922	Arabidops
c	55	45	60.8	174668	9	AC144848	Mus muscu
c	56	45	60.8	214567	14	AC084316	Mus muscu
c	57	45	60.8	227834	4	AC150561	Bos tauru
c	58	45	60.8	232433	9	AC153793	Mus muscu
c	59	45	59.5	1256	1	AF005700	Citrobact
c	60	44	59.5	3452	4	AF064209	Macropus
c	61	44	59.5	100000	8	AB020859	Homo sapi
c	62	44	59.5	110000	15	CR382126	Continuation (7 of
c	63	44	59.5	110000	15	AP008209	Continuation (129
c	64	44	59.5	110000	15	AP008209	Continuation (130
c	65	44	59.5	110000	15	AP008209	Continuation (130
c	66	44	59.5	118075	15	AC118675	Genomic s
c	67	44	59.5	129423	8	AC006362	Homo sapi
c	68	44	59.5	154701	14	AC120492	Rattus no
c	69	44	59.5	171914	5	AL929264	Zebrafish
c	70	44	59.5	191288	14	AC128123	Rattus no
c	71	44	59.5	204424	14	AC128159	Rattus no
c	72	44	59.5	204466	9	AC113463	Rattus no
c	73	44	59.5	205870	14	AC106907	Mus muscu
c	74	44	59.5	219782	14	AC114211	Rattus no
c	75	44	59.5	220069	9	AL732470	Mouse DNA
c	76	44	59.5	224784	14	CR450685	Danio rer

77	44	59.5	238934	9	AC084073	150	42	56.8	27024	15	AC139172	AC139172 Oryza sat
78	44	59.5	253649	14	AC106142	151	42	56.8	34794	2	UI3071	UI3071 Caenorhabdi
79	44	59.5	280257	14	AC096983	c 152	42	56.8	34836	8	AC006522	AC006522 Homo sapi
80	43.5	58.8	110000	1	CP000058.38	153	42	56.8	38675	8	AC015552	AC015552 Homo sapi
81	42	58.1	210	6	AR553628	154	42	56.8	42441	14	AC109339	AC109339 Homo sapi
82	43	58.1	474	13	AB062049	155	42	56.8	69128	14	AC017764	AC017764 Drosophil
83	43	58.1	474	13	AB062050	156	42	56.8	71736	15	AB009048	AB009048 Arabidops
84	43	58.1	474	13	AB062051	c 157	42	56.8	74832	15	AB016877	AB016877 Arabidops
85	43	58.1	474	13	AB062052	158	42	56.8	78736	5	BX548075	BX548075 Zebrafish
86	43	58.1	474	13	AB062053	159	42	56.8	96107	14	CR382324	CR382324 Arabidops
87	43	58.1	474	13	AB062054	c 160	42	56.8	104807	15	AC096605	AC096605 Arabidops
88	43	58.1	474	13	AB062055	c 161	42	56.8	110000	1	AB017282	AB017282 Oryza sat
89	43	58.1	474	13	AB062056	c 162	42	56.8	110000	14	AC095209	AC095209 Oryza sat
90	43	58.1	474	13	AB062057	c 163	42	56.8	110000	14	AC095209	AC095209 Oryza sat
91	43	58.1	474	13	AB062058	c 164	42	56.8	110000	14	AC095209	AC095209 Oryza sat
92	43	58.1	474	13	AB062059	c 165	42	56.8	110000	14	AC095209	AC095209 Oryza sat
93	43	58.1	474	13	AB062060	c 166	42	56.8	110000	14	AC095209	AC095209 Oryza sat
94	43	58.1	474	13	AB062061	c 167	42	56.8	110000	14	AC095209	AC095209 Oryza sat
95	43	58.1	474	13	AB062062	c 168	42	56.8	110000	14	AC095209	AC095209 Oryza sat
96	43	58.1	474	13	AB062063	c 169	42	56.8	110000	14	AC095209	AC095209 Oryza sat
97	43	58.1	474	13	AB062064	c 170	42	56.8	110000	14	AC095209	AC095209 Oryza sat
98	43	58.1	474	13	AB062065	c 171	42	56.8	110000	14	AC095209	AC095209 Oryza sat
99	43	58.1	474	13	AB062066	c 172	42	56.8	110000	14	AC095209	AC095209 Oryza sat
100	43	58.1	474	13	AB062067	c 173	42	56.8	110000	14	AC095209	AC095209 Oryza sat
101	43	58.1	474	13	AB062068	c 174	42	56.8	110000	14	AC095209	AC095209 Oryza sat
102	43	58.1	474	13	AB062069	c 175	42	56.8	110000	14	AC095209	AC095209 Oryza sat
103	43	58.1	474	13	AB062070	c 176	42	56.8	110000	14	AC095209	AC095209 Oryza sat
104	43	58.1	474	13	AB062071	c 177	42	56.8	110000	14	AC095209	AC095209 Oryza sat
105	43	58.1	474	13	AB062072	c 178	42	56.8	110000	14	AC095209	AC095209 Oryza sat
106	43	58.1	474	13	AB062073	c 179	42	56.8	110000	14	AC095209	AC095209 Oryza sat
107	43	58.1	474	13	AB062074	c 180	42	56.8	110000	14	AC095209	AC095209 Oryza sat
108	43	58.1	474	13	AB062075	c 181	42	56.8	110000	14	AC095209	AC095209 Oryza sat
109	43	58.1	474	13	AB062076	c 182	42	56.8	110000	14	AC095209	AC095209 Oryza sat
110	43	58.1	474	13	AB062077	c 183	42	56.8	110000	14	AC095209	AC095209 Oryza sat
111	43	58.1	474	13	AB062078	c 184	42	56.8	110000	14	AC095209	AC095209 Oryza sat
112	43	58.1	474	13	AB062079	c 185	42	56.8	110000	14	AC095209	AC095209 Oryza sat
113	43	58.1	474	13	AB062080	c 186	42	56.8	110000	14	AC095209	AC095209 Oryza sat
114	43	58.1	474	13	AB062081	c 187	42	56.8	110000	14	AC095209	AC095209 Oryza sat
115	43	58.1	474	13	AB062082	c 188	42	56.8	110000	14	AC095209	AC095209 Oryza sat
116	43	58.1	474	13	AB062083	c 189	42	56.8	110000	14	AC095209	AC095209 Oryza sat
117	43	58.1	474	13	AB062084	c 190	42	56.8	110000	14	AC095209	AC095209 Oryza sat
118	43	58.1	474	13	AB062085	c 191	42	56.8	110000	14	AC095209	AC095209 Oryza sat
119	43	58.1	474	13	AB062086	c 192	42	56.8	110000	14	AC095209	AC095209 Oryza sat
120	43	58.1	474	13	AB062087	c 193	42	56.8	110000	14	AC095209	AC095209 Oryza sat
121	43	58.1	474	13	AB062088	c 194	42	56.8	110000	14	AC095209	AC095209 Oryza sat
122	43	58.1	474	13	AB062089	c 195	42	56.8	110000	14	AC095209	AC095209 Oryza sat
123	43	58.1	474	13	AB062090	c 196	42	56.8	110000	14	AC095209	AC095209 Oryza sat
124	43	58.1	474	13	AB062091	c 197	42	56.8	110000	14	AC095209	AC095209 Oryza sat
125	43	58.1	474	13	AB062092	c 198	42	56.8	110000	14	AC095209	AC095209 Oryza sat
126	43	58.1	474	13	AB062093	c 199	42	56.8	110000	14	AC095209	AC095209 Oryza sat
127	43	58.1	474	13	AB062094	c 200	42	56.8	110000	14	AC095209	AC095209 Oryza sat
128	43	58.1	474	13	AB062095	c 201	42	56.8	110000	14	AC095209	AC095209 Oryza sat
129	43	58.1	474	13	AB062096	c 202	42	56.8	110000	14	AC095209	AC095209 Oryza sat
130	43	58.1	474	13	AB062097	c 203	42	56.8	110000	14	AC095209	AC095209 Oryza sat
131	43	58.1	474	13	AB062098	c 204	42	56.8	110000	14	AC095209	AC095209 Oryza sat
132	43	58.1	474	13	AB062099	c 205	42	56.8	110000	14	AC095209	AC095209 Oryza sat
133	43	58.1	474	13	AB062100	c 206	42	56.8	110000	14	AC095209	AC095209 Oryza sat
134	43	58.1	474	13	AB062101	c 207	42	56.8	110000	14	AC095209	AC095209 Oryza sat
135	43	58.1	474	13	AB062102	c 208	42	56.8	110000	14	AC095209	AC095209 Oryza sat
136	43	58.1	474	13	AB062103	c 209	42	56.8	110000	14	AC095209	AC095209 Oryza sat
137	43	58.1	474	13	AB062104	c 210	42	56.8	110000	14	AC095209	AC095209 Oryza sat
138	43	58.1	474	13	AB062105	c 211	42	56.8	110000	14	AC095209	AC095209 Oryza sat
139	43	58.1	474	13	AB062106	c 212	42	56.8	110000	14	AC095209	AC095209 Oryza sat
140	43	58.1	474	13	AB062107	c 213	42	56.8	110000	14	AC095209	AC095209 Oryza sat
141	43	58.1	474	13	AB062108	c 214	42	56.8	110000	14	AC095209	AC095209 Oryza sat
142	43	58.1	474	13	AB062109	c 215	42	56.8	110000	14	AC095209	AC095209 Oryza sat
143	43	58.1	474	13	AB062110	c 216	42	56.8	110000	14	AC095209	AC095209 Oryza sat
144	43	58.1	474	13	AB062111	c 217	42	56.8	110000	14	AC095209	AC095209 Oryza sat
145	43	58.1	474	13	AB062112	c 218	42	56.8	110000	14	AC095209	AC095209 Oryza sat
146	43	58.1	474	13	AB062113	c 219	42	56.8	110000	14	AC095209	AC095209 Oryza sat
147	43	58.1	474	13	AB062114	c 220	42	56.8	110000	14	AC095209	AC095209 Oryza sat
148	43	58.1	474	13	AB062115	c 221	42	56.8	110000	14	AC095209	AC095209 Oryza sat
149	43	58.1	474	13	AB062116	c 222	42	56.8	110000	14	AC095209	AC095209 Oryza sat

C 369	41	55.4	192944	14	AC146902	AC146902 Callicebu	442	41	55.4	271321	14	AC160174	AC160174 Bos tauru
C 370	41	55.4	193090	14	AC130049	AC130049 Rattus no	443	41	55.4	272590	14	AC152643	AC152643 Bos tauru
C 371	41	55.4	194001	14	AC151251	AC151251 Bos tauru	C 444	41	55.4	272590	14	AC152643	AC152643 Bos tauru
C 372	41	55.4	194006	8	AC074044	AC074044 Homo sapi	C 445	41	55.4	273962	14	AC160157	AC160157 Bos tauru
C 373	41	55.4	194826	14	AC163791	AC163791 Bos tauru	C 446	41	55.4	274760	14	AC156043	AC156043 Bos tauru
C 374	41	55.4	195512	9	AC133096	AC133096 Mus muscu	C 447	41	55.4	279137	14	AC152965	AC152965 Bos tauru
C 375	41	55.4	195778	5	BX088566	BX088566 Zebrafish	C 448	41	55.4	281195	14	AC163124	AC163124 Bos tauru
C 376	41	55.4	197063	14	AC159724	AC159724 Bos tauru	C 449	41	55.4	283651	14	AC153396	AC153396 Bos tauru
C 377	41	55.4	198450	14	AC149880	AC149880 Xenopus t	C 450	41	55.4	290052	14	AC160578	AC160578 Bos tauru
C 378	41	55.4	198669	15	ATCHRIV78	AL161582 Arabidops	C 451	41	55.4	290628	1	AE016983	AE016983 Shigella
C 379	41	55.4	198916	14	AC145008	AC145008 Bos tauru	C 452	41	55.4	291738	14	AC097744	AC097744 Rattus no
C 380	41	55.4	199111	14	AC158083	AC158083 Bos tauru	C 453	41	55.4	296706	14	CR759751	CR759751 Danio rer
C 381	41	55.4	202103	14	AC121195	AC121195 Rattus no	C 454	41	55.4	297936	14	AC156666	AC156666 Bos tauru
C 382	41	55.4	202353	14	AC164064	AC164064 Bos tauru	C 455	41	55.4	298590	14	AC159087	AC159087 Bos tauru
C 383	41	55.4	202759	14	AC152761	AC152761 Bos tauru	C 456	41	55.4	298896	14	AC151081	AC151081 Bos tauru
C 384	41	55.4	204091	14	AC152442	AC152442 Bos tauru	C 457	41	55.4	300029	15	AE017055	AE017055 Oryza sat
C 385	41	55.4	204214	9	AL672251	AL672251 Mouse DNA	C 458	41	55.4	302171	1	AE017295	AE017295 Leptospir
C 386	41	55.4	205028	9	AC154862	AC154862 Bos tauru	C 459	41	55.4	314531	14	CR388228	CR388228 Danio rer
C 387	41	55.4	205114	14	AC162489	AC162489 Bos tauru	C 460	41	55.4	315743	14	AC162474	AC162474 Bos tauru
C 388	41	55.4	205371	14	AC157157	AC157157 Bos tauru	C 461	41	55.4	316511	14	AC159369	AC159369 Bos tauru
C 389	41	55.4	207506	14	AC164053	AC164053 Bos tauru	C 462	41	55.4	320841	14	AC160006	AC160006 Bos tauru
C 390	41	55.4	208846	14	AC162339	AC162339 Bos tauru	C 463	41	55.4	324542	14	AC120727	AC120727 Rattus no
C 391	41	55.4	213195	14	AC156422	AC156422 Bos tauru	C 464	41	55.4	349652	1	BX571866	BX571866 Photorhab
C 392	41	55.4	213380	14	AC156507	AC156507 Bos tauru	C 465	41	55.4	349980	6	AX770906	AX770906 Sequence
C 393	41	55.4	214259	14	AC160786	AC160786 Bos tauru	C 466	40.5	54.7	1262	1	AF274326	AF274326 Pseudomon
C 394	41	55.4	215374	14	AC150765	AC150765 Bos tauru	C 467	40.5	54.7	78855	14	AC166084	AC166084 Bos tauru
C 395	41	55.4	216226	14	AC137215	AC137215 Rattus no	C 468	40.5	54.7	80169	14	AC164773	AC164773 Bos tauru
C 396	41	55.4	219466	14	AC164018	AC164018 Bos tauru	C 469	40	54.1	165	6	CQ751328	CQ751328 Sequence
C 397	41	55.4	220640	14	AC162093	AC162093 Bos tauru	C 470	40	54.1	326	13	AB113308	AB113308 Hepatitis
C 398	41	55.4	220714	14	CT010454	CT010454 Mus muscu	C 471	40	54.1	326	13	AB200243	AB200243 Hepatitis
C 399	41	55.4	221566	14	AC161951	AC161951 Bos tauru	C 472	40	54.1	430	2	AF300460	AF300460 Axinella
C 400	41	55.4	223859	14	AC159051	AC159051 Bos tauru	C 473	40	54.1	459	6	CQ173785	CQ173785 Sequence
C 401	41	55.4	225884	14	AC098500	AC098500 Rattus no	C 474	40	54.1	524	10	BV002022	BV002022 S209P6145
C 402	41	55.4	227026	9	AC113476	AC113476 Mus muscu	C 475	40	54.1	567	6	CQ178166	CQ178166 Sequence
C 403	41	55.4	227119	14	BX119971	BX119971 Danio rer	C 476	40	54.1	628	10	BV328837	BV328837 S241P6199
C 404	41	55.4	227199	14	AC094896	AC094896 Rattus no	C 477	40	54.1	672	10	BV517386	BV517386 r1t18b10.
C 405	41	55.4	227278	14	AC149654	AC149654 Bos tauru	C 478	40	54.1	708	10	BV538201	BV538201 G591P6358
C 406	41	55.4	228030	14	AC164232	AC164232 Bos tauru	C 479	40	54.1	718	8	HUMDTN21	U84549 Human dystr
C 407	41	55.4	228079	14	AC156851	AC156851 Bos tauru	C 480	40	54.1	725	10	BV541075	BV541075 G591P5139
C 408	41	55.4	228307	14	AC153699	AC153699 Bos tauru	C 481	40	54.1	743	10	BV646733	BV646733 S217P6032
C 409	41	55.4	228864	14	AC155755	AC155755 Bos tauru	C 482	40	54.1	769	8	HS4342097	HS4342097 Homo sapi
C 410	41	55.4	229327	14	AC160441	AC160441 Bos tauru	C 483	40	54.1	795	10	BV542081	BV542081 Mycoplasma
C 411	41	55.4	229396	14	AC024108	AC024108 Homo sapi	C 484	40	54.1	878	10	BV547173	BV547173 rrj65c06.
C 412	41	55.4	231452	14	AC160482	AC160482 Bos tauru	C 485	40	54.1	928	10	BV520708	BV520708 Mycoplasma
C 413	41	55.4	232542	14	AC159393	AC159393 Bos tauru	C 486	40	54.1	955	1	AF314228	AF314228 Mycoplasma
C 414	41	55.4	232726	14	AC159787	AC159787 Bos tauru	C 487	40	54.1	1055	1	AF143444	AF143444 Sinorhizo
C 415	41	55.4	232737	2	AE003660	AE003660 Drosophil	C 488	40	54.1	1125	6	CQ733313	CQ733313 Sequence
C 416	41	55.4	232952	14	AC157126	AC157126 Bos tauru	C 489	40	54.1	1182	8	AB060194	AB060194 Macaca fa
C 417	41	55.4	234856	14	AC157006	AC157006 Bos tauru	C 490	40	54.1	1468	15	YLSRPP19A	YLSRPP19A
C 418	41	55.4	236385	14	AC110349	AC110349 Rattus no	C 491	40	54.1	1801	8	HSU46746	HSU46746 Human dystr
C 419	41	55.4	237065	14	AC128999	AC128999 Rattus no	C 492	40	54.1	2080	5	BC090464	BC090464 Danio rer
C 420	41	55.4	237450	14	AC095092	AC095092 Rattus no	C 493	40	54.1	2194	1	SME244574	SM244574 Sinorhizo
C 421	41	55.4	237986	14	AC106932	AC106932 Rattus no	C 494	40	54.1	2226	1	AF488710	AF488710 Mycoplasma
C 422	41	55.4	238712	14	AC115308	AC115308 Rattus no	C 495	40	54.1	2310	6	CQ817586	CQ817586 Sequence
C 423	41	55.4	239760	14	AC095113	AC095113 Rattus no	C 496	40	54.1	2518	8	HSU46744	HSU46744 Human dystr
C 424	41	55.4	239905	14	AC154963	AC154963 Bos tauru	C 497	40	54.1	2591	6	AX833370	AX833370 Sequence
C 425	41	55.4	243486	14	AC153262	AC153262 Bos tauru	C 498	40	54.1	2991	8	AK095037	AK095037 Homo sapi
C 426	41	55.4	244183	14	AC160787	AC160787 Bos tauru	C 499	40	54.1	2695	4	AF415201S1	AF415201 Sue scrof
C 427	41	55.4	244429	14	AC155061	AC155061 Bos tauru	C 500	40	54.1	3272	15	AY255705	AY255705 Mangifera
C 428	41	55.4	248040	14	AC163839	AC163839 Bos tauru							
C 429	41	55.4	248151	14	AC157338	AC157338 Bos tauru							
C 430	41	55.4	249153	14	AC162990	AC162990 Bos tauru							
C 431	41	55.4	250898	14	AC153042	AC153042 Bos tauru							
C 432	41	55.4	252650	14	AC160208	AC160208 Bos tauru							
C 433	41	55.4	255140	14	AC160814	AC160814 Bos tauru							
C 434	41	55.4	256265	14	AC122943	AC122943 Rattus no							
C 435	41	55.4	256530	14	AC109032	AC109032 Rattus no							
C 436	41	55.4	257145	14	AC163794	AC163794 Bos tauru							
C 437	41	55.4	258016	14	AC163323	AC163323 Bos tauru							
C 438	41	55.4	258031	14	AC163117	AC163117 Bos tauru							
C 439	41	55.4	261408	14	BX950188	BX950188 Danio rer							
C 440	41	55.4	266711	14	AC163472	AC163472 Bos tauru							
C 441	41	55.4	269573	14	AC162560	AC162560 Bos tauru							

ALIGNMENTS

RESULT 1	BD211562	345 bp	DNA	linear	PAT 17-JUL-2003
BD211562	Canine and feline immunoregulatory proteins, nucleic acid molecules				
LOCUS	Canine and feline immunoregulatory proteins, nucleic acid molecules				
DEFINITION	Canine and feline immunoregulatory proteins, nucleic acid molecules				
ACCESSION	BD211562				
VERSION	BD211562.1				
KEYWORDS	JP 2002516104-A/68.				
SOURCE	Canis familiaris (dog)				
ORGANISM	Canis familiaris				

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae; Canis.

REFERENCE 1 (bases 1 to 345)
 AUTHORS Sim, G., Yang S., Dreitz, M.J. and Wonderling, R.S.
 TITLE Canine and feline immunoregulatory proteins, nucleic acid molecules and method of using the same
 JOURNAL Patent: JP 2002516104-A 68 04-JUN-2002;

HESKA CORP

COMMENT OS Canis familiaris (dog)
 PN JP 2002516104-A/68
 PD 04-JUN-2002
 PF 28-MAY-1999 JP 2000551002
 PR 29-MAY-1998 US 60/087306

PI GEKKEE SIM, SHUMIN YANG, MATTHEW J DREITZ, RAMANI S WONDERLING PC
 C12N15/09, A61K31/7088, A61K38/00, A61K38/21, A61K39/00, A61K39/395,
 PC A61K39/395.

PC A61K45/00, A61K48/00, A61P37/02, A61P37/04, C07K14/475, C07K14/535,
 PC C07K14/54,
 PC C07K14/56, C07K14/705, C07K16/24, C07K16/28, C12N1/21, C12N5/10, PC
 G01N33/15,

PC G01N33/50, C12N15/00, A61K37/02, A61K37/66, C12N5/00 CC Canine
 and feline immunoregulatory proteins, nucleic acid CC

molecules and
 CC method of using the same
 FH Key Location/Qualifiers
 FT CDS (1)..(345).

FEATURES
 source
 1..345
 Location/Qualifiers
 /organism="Canis familiaris"
 /mol_type="genomic DNA"
 /db_xref="taxon:9615"

ORIGIN

Alignment Scores:
 Pred. No.: 3.57e-06 Length: 345
 Score: 74.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-787-382-20 (1-15) x BD211562 (1-345)

Qy 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
 |||||
 Db 1 TTTCGTGTAGAAAATCCCATGATAGACTGGTGGCAGACCTTG 45

RESULT 2
 BD211563/c

LOCUS BD211563 345 bp DNA linear PAT 17-JUL-2003
 DEFINITION Canine and feline immunoregulatory proteins, nucleic acid molecules and method of using the same.

ACCESSION BD211563
 VERSION BD211563.1 GI:33021333
 KEYWORDS JP 2002516104-A/69.
 SOURCE Canis familiaris (dog)
 ORGANISM Canis familiaris

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae; Canis.

REFERENCE 1 (bases 1 to 345)
 AUTHORS Sim, G., Yang S., Dreitz, M.J. and Wonderling, R.S.
 TITLE Canine and feline immunoregulatory proteins, nucleic acid molecules and method of using the same
 JOURNAL Patent: JP 2002516104-A 69 04-JUN-2002;

HESKA CORP

COMMENT OS Canis familiaris (dog)
 PN JP 2002516104-A/69
 PD 04-JUN-2002
 PF 28-MAY-1999 JP 2000551002
 PR 29-MAY-1998 US 60/087306

PI GEKKEE SIM, SHUMIN YANG, MATTHEW J DREITZ, RAMANI S WONDERLING PC

C12N15/09, A61K31/7088, A61K38/00, A61K38/21, A61K39/00, A61K39/395,
 PC A61K39/395,
 PC A61K45/00, A61K48/00, A61P37/02, A61P37/04, C07K14/475, C07K14/535,
 PC C07K14/54,
 PC C07K14/56, C07K14/705, C07K16/24, C07K16/28, C12N1/21, C12N5/10, PC
 G01N33/15,
 PC G01N33/50, C12N15/00, A61K37/02, A61K37/66, C12N5/00 CC Canine
 and feline immunoregulatory proteins, nucleic acid CC

molecules and
 CC method of using the same
 FH Key Location/Qualifiers
 FT source 1..345
 /organism="Canis familiaris (dog)".

FEATURES
 source
 1..345
 Location/Qualifiers
 /organism="Canis familiaris"
 /mol_type="genomic DNA"
 /db_xref="taxon:9615"

ORIGIN

Alignment Scores:
 Pred. No.: 3.57e-06 Length: 345
 Score: 74.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-787-382-20 (1-15) x BD211563 (1-345)

Qy 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
 |||||
 Db 345 TTTCGTGTAGAAAATCCCATGATAGACTGGTGGCAGACCTTG 301

RESULT 3
 AR241540

LOCUS AR241540 345 bp DNA linear PAT 20-DEC-2002
 DEFINITION Sequence 85 from patent US 6471957.
 ACCESSION AR241540
 VERSION AR241540.1 GI:27287249
 KEYWORDS Unknown.
 SOURCE Unknown.
 ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 345)
 AUTHORS Sim, G.-K., Yang S., Dreitz, M.J. and Wonderling, R.S.
 TITLE Canine IL-4 immunoregulatory proteins and uses thereof
 JOURNAL Patent: US 6471957-A 85 29-OCT-2002;
 HESKA Corporation; Fort Collins, CO;
 EPX;

FEATURES
 source
 1..345
 Location/Qualifiers
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN

Alignment Scores:
 Pred. No.: 3.57e-06 Length: 345
 Score: 74.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-787-382-20 (1-15) x AR241540 (1-345)

Qy 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
 |||||
 Db 1 TTTCGTGTAGAAAATCCCATGATAGACTGGTGGCAGACCTTG 45

RESULT 4
 AR241541/c

LOCUS AR241541 345 bp DNA linear PAT 20-DEC-2002

```

DEFINITION Sequence 87 from patent US 6471957.
ACCESSION AR241541
VERSION AR241541.1 GI:27287250
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
  1 (bases 1 to 345)
  Unclassified.
AUTHORS Sim,G.-K., Yang,S., Dreibitz,M.J. and Wonderling,R.S.
TITLE Canine IL-4 immunoregulatory proteins and uses thereof
JOURNAL Patent: US 6471957-A 87 29-OCT-2002;
  Heska Corporation; Fort Collins, CO;
  EPX;
FEATURES
  source
    Location/Qualifiers
    1..345
    /organism="unknown"
    /mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 345 Length: 345
Score: 74.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-787-382-20 (1-15) x AR241541 (1-345)
QY 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
Db 345 TTTGCTGTAGAAATCCCATGATAGACTGGTGGCAGACCTTG 301
RESULT 5
AR254496
LOCUS AR254496 345 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 85 from patent US 6482403.
ACCESSION AR254496
VERSION AR254496.1 GI:27303384
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
  1 (bases 1 to 345)
  Unclassified.
AUTHORS Sim,G.-K., Yang,S., Dreibitz,M.J. and Wonderling,R.S.
TITLE Canine IL-13 immunoregulatory proteins and uses thereof
JOURNAL Patent: US 6482403-A 85 19-NOV-2002;
  Heska Corporation; Fort Collins, CO
FEATURES
  source
    Location/Qualifiers
    1..345
    /organism="unknown"
    /mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 345 Length: 345
Score: 74.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-787-382-20 (1-15) x AR254496 (1-345)
QY 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
Db 1 TTTGCTGTAGAAATCCCATGATAGACTGGTGGCAGACCTTG 45
RESULT 6
AR254497/c
LOCUS AR254497 345 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 87 from patent US 6482403.
ACCESSION AR254497
VERSION AR254497.1 GI:27303385
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
  1 (bases 1 to 345)
  Unclassified.
AUTHORS Sim,G.-K., Yang,S., Dreibitz,M.J. and Wonderling,R.S.
TITLE Canine IL-13 immunoregulatory proteins and uses thereof
JOURNAL Patent: US 6482403-A 87 19-NOV-2002;
  Heska Corporation; Fort Collins, CO
FEATURES
  source
    Location/Qualifiers
    1..345
    /organism="unknown"
    /mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 345 Length: 345
Score: 74.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-787-382-20 (1-15) x AR254497 (1-345)
QY 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
Db 345 TTTGCTGTAGAAATCCCATGATAGACTGGTGGCAGACCTTG 301
RESULT 7
AR254497
LOCUS AR254497 356 bp mRNA linear MAM 04-AUG-1999
DEFINITION Canis familiaris interleukin-5 mRNA, partial cds.
ACCESSION AF091133
VERSION AF091133.1 GI:5690203
KEYWORDS
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris
REFERENCE
  1 (bases 1 to 356)
  German,A.J., Helps,C.R., Harley,R., Hall,E.J. and Day,M.J.
  Cloning and sequencing of canine interleukin-5
  Unpublished
  2 (bases 1 to 356)
  German,A.J., Helps,C.R., Harley,R., Hall,E.J. and Day,M.J.
  Direct Submission
  Submitted (12-SEP-1998) Department of Clinical Veterinary Science,
  University of Bristol, Langford House, Langford, Bristol, North
  Somerset BS40 5DU, United Kingdom
  Location/Qualifiers
  1..356
  /organism="Canis familiaris"
  /mol_type="mRNA"
  /db_xref="taxon:9615"
  <!.>356
  /note="IL-5"
  /codon_start=1
  /product="interleukin-5"
  /protein_id="AAD46991.1"
  /db_xref="GI:5690204"
  /translation="LGVLCSAFAVENPMNRLVAETLTLLSTHRTWLTIGDNLMIPTPE
  NKNHQCIKEVFGIDITLKNQTARGEAVDKLFQNLSLIKEHIEROKRCAGKKKVKVTN
  FLNVLQLFLGVNMTW"
ORIGIN
Alignment Scores:
Pred. No.: 345 Length: 356
Score: 74.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

```

```

Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0
US-10-787-382-20 (1-15) x AF091133 (1-356)
QY 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
DB 22 TTTCGCTGAGAAATCCCATGATAGACTGGTGGCAGACCTTG 66

RESULT 8
BD211560
LOCUS BD211560 402 bp DNA linear PAT 17-JUL-2003
DEFINITION Canine and feline immunoregulatory proteins, nucleic acid molecules
and method of using the same.
ACCESSION BD211560
VERSION BD211560.1 GI:33021330
KEYWORDS JP 2002516104-A/66.
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
Canis.
REFERENCE 1 (bases 1 to 402)
AUTHORS Sim,G., Yang,S., Dreitz,M.J. and Wonderling,R.S.
TITLE Canine and feline immunoregulatory proteins, nucleic acid molecules
and method of using the same
JOURNAL Patent: JP 2002516104-A 66 04-JUN-2002;
COMMENT HESKA CORP
OS Canis familiaris (dog)
PN JP 2002516104-A/66
PD 04-JUN-2002
PF 28-MAY-1999 JP 2000551002
PR 29-MAY-1998 US 60/087306
PI GEKKEE SIM, SHUMIN YANG, MATTHEW J DREITZ, RAMANI S WONDERLING PC
C12N15/09, A61K31/7088, A61K38/00, A61K39/00, A61K39/395,
PC A61K39/395,
PC A61K45/00, A61K48/00, A61P37/02, A61P37/04, C07K14/475, C07K14/535,
PC C07K14/54,
PC C07K14/56, C07K14/705, C07K16/24, C07K16/28, C12N1/21, C12N5/10, PC
G01N33/15,
PC G01N33/50, C12N15/00, A61K37/02, A61K37/66, C12N5/00 CC Canine
and feline immunoregulatory proteins, nucleic acid CC
molecules and
CC method of using the same
FT Key Location/Qualifiers
FT source 1..402
FT /organism='Canis familiaris (dog)'
FEATURES
source
ORIGIN
Alignment Scores:
Pred. No.: 4.26e-06 Length: 402
Score: 74.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-787-382-20 (1-15) x BD211561 (1-402)
QY 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
DB 345 TTTCGCTGAGAAATCCCATGATAGACTGGTGGCAGACCTTG 301

RESULT 10
AR241538
LOCUS AR241538 402 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 83 from patent US 6471957.
ACCESSION AR241538
VERSION AR241538.1 GI:27287247
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 402)
AUTHORS Sim,G.-K., Yang,S., Dreitz,M.J. and Wonderling,R.S.
TITLE Canine IL-4 immunoregulatory proteins and uses thereof
JOURNAL Patent: US 6471957-A 83 29-OCT-2002;
Heska Corporation; Fort Collins, CO;
EPX,
FEATURES
source
ORIGIN
Alignment Scores:
Pred. No.: 4.26e-06 Length: 402
Score: 74.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-787-382-20 (1-15) x BD211560 (1-402)
QY 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
DB 58 TTTCGCTGAGAAATCCCATGATAGACTGGTGGCAGACCTTG 102

RESULT 9
BD211561/c
LOCUS BD211561/c 402 bp DNA linear PAT 17-JUL-2003
DEFINITION Canine and feline immunoregulatory proteins, nucleic acid molecules
and method of using the same.

```

/mol_type="genomic DNA"

ORIGIN

Alignment Scores:
 Pred. No.: 4.26e-06 Length: 402
 Score: 74.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-787-382-20 (1-15) x AR241538 (1-402)

QY 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrIeu 15
 |||||
 Db 58 TTTCGTGTAGAAAATCCCATGATAGACTGGTGGCAGACCTTG 102

RESULT 11
 AR241539/c
 LOCUS AR241539 402 bp DNA linear PAT 20-DEC-2002

DEFINITION Sequence 84 from patent US 6471957.
 ACCESSION AR241539
 VERSION AR241539.1 GI:27287248
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 402)
 AUTHORS Sim,G.-K., Yang,S., Dreitz,M.J. and Wonderling,R.S.
 TITLE Canine IL-4 immunoregulatory proteins and uses thereof
 JOURNAL Patent: US 6471957-A 84 29-OCT-2002;
 Heska Corporation; Fort Collins, CO;
 EPX;

FEATURES
 source
 1..402
 Location/Qualifiers
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN

Alignment Scores:
 Pred. No.: 4.26e-06 Length: 402
 Score: 74.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-787-382-20 (1-15) x AR241539 (1-402)

QY 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrIeu 15
 |||||
 Db 345 TTTCGTGTAGAAAATCCCATGATAGACTGGTGGCAGACCTTG 301

RESULT 12
 AR254494
 LOCUS AR254494 402 bp DNA linear PAT 20-DEC-2002

DEFINITION Sequence 83 from patent US 6482403.
 ACCESSION AR254494
 VERSION AR254494.1 GI:27303382
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 402)
 AUTHORS Sim,G.-K., Yang,S., Dreitz,M.J. and Wonderling,R.S.
 TITLE Canine IL-13 immunoregulatory proteins and uses thereof
 JOURNAL Patent: US 6482403-A 83 19-NOV-2002;
 Heska Corporation; Fort Collins, CO

FEATURES
 source
 1..402
 Location/Qualifiers
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN

Alignment Scores:
 Pred. No.: 4.26e-06 Length: 402
 Score: 74.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-787-382-20 (1-15) x AR254494 (1-402)

QY 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrIeu 15
 |||||
 Db 58 TTTCGTGTAGAAAATCCCATGATAGACTGGTGGCAGACCTTG 102

RESULT 13
 AR254495/c

LOCUS AR254495 402 bp DNA linear PAT 20-DEC-2002

DEFINITION Sequence 84 from patent US 6482403.
 ACCESSION AR254495
 VERSION AR254495.1 GI:27303383
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 402)
 AUTHORS Sim,G.-K., Yang,S., Dreitz,M.J. and Wonderling,R.S.
 TITLE Canine IL-13 immunoregulatory proteins and uses thereof
 JOURNAL Patent: US 6482403-A 84 19-NOV-2002;
 Heska Corporation; Fort Collins, CO
 FEATURES
 Location/Qualifiers
 source
 1..402
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN

Alignment Scores:
 Pred. No.: 4.26e-06 Length: 402
 Score: 74.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-787-382-20 (1-15) x AR254495 (1-402)

QY 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrIeu 15
 |||||
 Db 345 TTTCGTGTAGAAAATCCCATGATAGACTGGTGGCAGACCTTG 301

RESULT 14
 AR300436

LOCUS AR300436 405 bp DNA linear PAT 12-JUN-2003

DEFINITION Sequence 1 from patent US 6537781.
 ACCESSION AR300436
 VERSION AR300436.1 GI:31687875
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 405)
 AUTHORS Guo,H., Lawton,R., Mermer,B. and Aliyappa,A.P.
 TITLE Methods and compositions concerning canine interleukin 5
 JOURNAL Patent: US 6537781-A 1 25-MAR-2003;
 IDEXX Laboratories, Inc.; Westbrook, ME

FEATURES
 source
 1..405
 Location/Qualifiers
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN

Alignment Scores:
 Pred. No.: 4.3e-06 Length: 405

```

Score: 74.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-787-382-20 (1-15) x AR300436 (1-405)

Qy 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
|||||
Db 58 TTTCGTGTAGAAATCCCATGATAGACTGTGGCAGACCTTG 102

RESULT 15
LOCUS AF331919 610 bp DNA linear PAT 22-JUN-2001
DEFINITION Canis familiaris interleukin-5 mRNA, complete cds.
ACCESSION AF331919
VERSION AF331919.1 GI:15919180
KEYWORDS Canis familiaris (dog)
SOURCE Canis familiaris
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
Canis.
REFERENCE 1
AUTHORS Guo, H., Lawton, R., Mermer, B. and Aiyappa, A.P.
TITLE Methods and compositions concerning canine interleukin 5
JOURNAL IDEXX LABORATORIES, INC. (US)
COMMENT On Jun 24, 2001 this sequence version replaced gi:13185501.
FEATURES
source 1..405
/mol_type="unassigned DNA"
/db_xref="taxon:9615"

ORIGIN
Alignment Scores:
Pred. No.: 4.3e-06 Length: 405
Score: 74.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-787-382-20 (1-15) x AX083939 (1-405)

Qy 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
|||||
Db 58 TTTCGTGTAGAAATCCCATGATAGACTGTGGCAGACCTTG 102

RESULT 16
LOCUS AF331919 610 bp mRNA linear MAM 04-OCT-2001
DEFINITION Canis familiaris interleukin-5 mRNA, complete cds.
ACCESSION AF331919
VERSION AF331919.1 GI:15919180
KEYWORDS Canis familiaris (dog)
SOURCE Canis familiaris
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
Canis.
REFERENCE 1 (bases 1 to 610)
AUTHORS Yang, S., Sellins, K.S., Weber, E. and McCall, C.
TITLE Canine interleukin-5: molecular characterization of the gene and
expression of biologically active recombinant protein
JOURNAL J. Interferon Cytokine Res. 21 (6), 361-367 (2001)
PUBMED 11440633
REFERENCE 2 (bases 1 to 610)
AUTHORS Yang, S.
TITLE Direct Submission
JOURNAL Submitted (22-DEC-2000) Immunology, Heska Corporation, 1613

Prospect Parkway, Ft Collins, CO 80525, USA
Location/Qualifiers
1..610
/organism="Canis familiaris"
/mol_type="mRNA"
/db_xref="taxon:9615"
1..28
5'UTR
CDS
29..433
/note="IL-5"
/codon_start=1
/product="interleukin-5"
/protein_id="AAL10715.1"
/db_xref="GI:15919181"
/translation="MRMLNLNLSLALGAAYVSAFAVENPMRNLVAETLLSTHRTWL
IGDNLMIPTPENKHNQCICEVFGQIDTLKNQTAHGEAVDKLFQNLSLIKHIERQK
KRCAGERWRVTKFLDYLVQLVGLVINTWTPS"
433..610

3'UTR
ORIGIN

Alignment Scores:
Pred. No.: 6.91e-06 Length: 610
Score: 74.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-787-382-20 (1-15) x AF331919 (1-610)

Qy 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
|||||
Db 86 TTTCGTGTAGAAATCCCATGATAGACTGTGGCAGACCTTG 130

RESULT 17
LOCUS BD211558 610 bp DNA linear PAT 17-JUL-2003
DEFINITION Canine and feline immunoregulatory proteins, nucleic acid molecules
and method of using the same.
ACCESSION BD211558
VERSION BD211558.1 GI:33021328
KEYWORDS Canis familiaris (dog)
SOURCE Canis familiaris
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
Canis.
REFERENCE 1 (bases 1 to 610)
AUTHORS Sim, G., Yang, S., Dreitz, M.J. and Wonderling, R.S.
TITLE Canine and feline immunoregulatory proteins, nucleic acid molecules
and method of using the same
JOURNAL Patent: JP 2002516104-A 64 04-JUN-2002;
HESKA CORP
COMMENT OS Canis familiaris (dog)
PN JP 2002516104-A/64
PD 04-JUN-2002
PF 28-MAY-1999 JP 2000551002
PR 29-MAY-1998 US 60/087306
PI GEKKEE SIM, SHUMIN YANG, MATTHEW J DREITZ, RAMANI S WONDERLING PC
C12N15/09, A61K31/7088, A61K39/00, A61K38/21, A61K39/00, A61K39/395,
PC A61K39/395,
PC A61K45/00, A61K48/00, A61P37/02, A61P37/04, C07K14/475, C07K14/535,
PC C07K14/54,
PC C07K14/56, C07K14/705, C07K16/24, C07K16/28, C12N1/21, C12N5/10, PC
G01N33/15,
PC G01N33/50, C12N15/00, A61K37/02, A61K37/66, C12N5/00 CC Canine
and feline immunoregulatory proteins, nucleic acid CC
molecules and
CC method of using the same
FH Key Location/Qualifiers
FT CDS (29)..(430).
Location/Qualifiers
1..610
/organism="Canis familiaris"

FEATURES
source

```

/mol_type="genomic DNA"
/db_xref="taxon:9615"

ORIGIN

Alignment Scores:
Pred. No.: 6.91e-06 Length: 610
Score: 74.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-787-382-20 (1-15) x BD211558 (1-610)

QY 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
|||||
DB 86 TTTGCTGTAGAAAATCCCATGAATAGACTGGTGGCAGACCTTG 130

RESULT 18

BD211559/c
LOCUS BD211559 610 bp DNA linear PAT 17-JUL-2003
DEFINITION Canine and feline immunoregulatory proteins, nucleic acid molecules
and method of using the same.

ACCESSION BD211559
VERSION BD211559.1 GI:33021329
KEYWORDS JP 2002516104-A/65.
SOURCE Canis familiaris (dog)

ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
Canis.

REFERENCE 1 (bases 1 to 610)

AUTHORS Sim,G., Yang,S., Dreibitz,M.J. and Wonderling,R.S.

TITLE Canine and feline immunoregulatory proteins, nucleic acid molecules
and method of using the same

JOURNAL Patent: JP 2002516104-A 65 04-JUN-2002;

COMMENT OS Canis familiaris (dog)
PN JP 2002516104-A/65
PD 04-JUN-2002

PF 28-MAY-1999 JP 2000551002

PI 29-MAY-1998 US 60/087306

PT GEKKEI SIM, SHUMIN YANG, MATTHEW J DREITZ, RAMANI S WONDERLING PC
C12N15/09, A61K31/7088, A61K38/00, A61K38/21, A61K39/00, A61K39/395,

PC A61K39/395,

PC A61K45/00, A61K48/00, A61P37/02, A61P37/04, C07K14/475, C07K14/535,

PC C07K14/54,

PC C07K14/56, C07K14/705, C07K16/24, C07K16/28, C12N1/21, C12N5/10, PC

G01N33/15,

PC G01N33/50, C12N15/00, A61K37/02, A61K37/66, C12N5/00 CC Canine
and feline immunoregulatory proteins, nucleic acid CC

molecules and

CC method of using the same

PH Key Location/Qualifiers

FT source 1. 610

FT Location/Qualifiers

/organism="Canis familiaris (dog)"

FEATURES

source

1. 610

/organism="Canis familiaris"

/mol_type="genomic DNA"

/db_xref="taxon:9615"

ORIGIN

Alignment Scores:
Pred. No.: 6.91e-06 Length: 610
Score: 74.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-787-382-20 (1-15) x BD211559 (1-610)

QY 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
|||||
DB 525 TTTGCTGTAGAAAATCCCATGAATAGACTGGTGGCAGACCTTG 481

RESULT 19

AR241536

LOCUS AR241536

DEFINITION Sequence 80 from patent US 6471957.

ACCESSION AR241536

VERSION AR241536.1 GI:27287245

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 610)

AUTHORS Sim,G.-K., Yang,S., Dreibitz,M.J. and Wonderling,R.S.

TITLE Canine IL-4 immunoregulatory proteins and uses thereof

JOURNAL Patent: US 6471957-A 80 29-OCT-2002;

Heska Corporation; Fort Collins, CO;

EPX;

FEATURES

source

1. 610

/organism="unknown"

/mol_type="genomic DNA"

ORIGIN

Alignment Scores:
Pred. No.: 6.91e-06 Length: 610
Score: 74.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-787-382-20 (1-15) x AR241536 (1-610)

QY 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
|||||

DB 86 TTTGCTGTAGAAAATCCCATGAATAGACTGGTGGCAGACCTTG 130

RESULT 20

AR241537/c

LOCUS AR241537

DEFINITION Sequence 82 from patent US 6471957.

ACCESSION AR241537

VERSION AR241537.1 GI:27287246

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 610)

AUTHORS Sim,G.-K., Yang,S., Dreibitz,M.J. and Wonderling,R.S.

TITLE Canine IL-4 immunoregulatory proteins and uses thereof

JOURNAL Patent: US 6471957-A 82 29-OCT-2002;

Heska Corporation; Fort Collins, CO;

EPX;

FEATURES

source

1. 610

/organism="unknown"

/mol_type="genomic DNA"

ORIGIN

Alignment Scores:
Pred. No.: 6.91e-06 Length: 610
Score: 74.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-787-382-20 (1-15) x AR241537 (1-610)

QY 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15


```
Db 525 TTTCGTAGAAAATCCCATGAATAGACTGGTGGCAGACCTTG 481
|||||
RESULT 21
AR254492
LOCUS AR254492 610 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 80 from patent US 6482403.
ACCESSION AR254492
VERSION AR254492.1 GI:27303380
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 610)
AUTHORS Sim,G.-K., Yang,S., Dreitz,M.J. and Wonderling,R.S.
TITLE Caniney IL-13 immunoregulatory proteins and uses thereof
JOURNAL Patent: US 6482403-A 80 19-NOV-2002;
Heska Corporation; Fort Collins, CO
FEATURES
source
1..610
/mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 6.91e-06 Length: 610
Score: 74.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-787-382-20 (1-15) x AR254492 (1-610)
Qy 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
Db 86 TTTCGTAGAAAATCCCATGAATAGACTGGTGGCAGACCTTG 130
|||||
RESULT 22
AR254493/c
LOCUS AR254493 610 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 82 from patent US 6482403.
ACCESSION AR254493
VERSION AR254493.1 GI:27303381
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 610)
AUTHORS Sim,G.-K., Yang,S., Dreitz,M.J. and Wonderling,R.S.
TITLE Caniney IL-13 immunoregulatory proteins and uses thereof
JOURNAL Patent: US 6482403-A 82 19-NOV-2002;
Heska Corporation; Fort Collins, CO
FEATURES
source
1..610
/mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 6.91e-06 Length: 610
Score: 74.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-787-382-20 (1-15) x AR254493 (1-610)
Qy 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
Db 525 TTTCGTAGAAAATCCCATGAATAGACTGGTGGCAGACCTTG 481
|||||
RESULT 23
AF331920
LOCUS AF331920 1658 bp DNA linear MAM 04-OCT-2001
DEFINITION Canis familiaris interleukin-5 gene, complete cds.
ACCESSION AF331920
VERSION AF331920.1 GI:15919182
KEYWORDS
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris
Canis familiaris
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
Canis.
REFERENCE 1 (bases 1 to 1658)
AUTHORS Yang,S., Sellins,K.S., Weber,E. and McCall,C.
TITLE Canine interon of interleukin-5: molecular characterization of the gene and
expression of biologically active recombinant protein
JOURNAL J Interferon Cytokine Res. 21 (6), 361-367 (2001)
PUBMED 11440833
REFERENCE 2 (bases 1 to 1658)
AUTHORS Yang,S.
TITLE Direct Submission
JOURNAL Submitted (22-DEC-2000) Immunology, Heska Corporation, 1613
Prospect Parkway, Ft Collins, CO 80525, USA
FEATURES
source
1..1658
/mol_type="genomic DNA"
/db_xref="taxon:9615"
join(1..170,374..406,1276..1404,1523..1658)
/mol_type="interleukin-5"
1..26
join(27..170,374..406,1276..1404,1523..1621)
/note="IL-5"
/codon_start=1
/product="interleukin-5"
/protein_id="AA110716.1"
/db_xref="GI:15919183"
/translation="MRMLNLNLALGAAYVSFAVENPMNRLVAETLTLLSTHRTWL
IGDGLMIPTEPNKHOLCIKEVFGIDTLKNQTAHGEAVDKLFQNLISLKEHIERQK
KRCAGERWRVTKFLDYLVQVFLGVINTEWTPES"
1622..1658
3'UTR
ORIGIN
Alignment Scores:
Pred. No.: 2.2e-05 Length: 1658
Score: 74.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0
US-10-787-382-20 (1-15) x AF331920 (1-1658)
Qy 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
Db 84 TTTCGTAGAAAATCCCATGAATAGACTGGTGGCAGACCTTG 128
|||||
RESULT 24
ECU91947
LOCUS ECU91947 405 bp mRNA linear MAM 05-APR-1997
DEFINITION Equus caballus interleukin-5 (IL-5) mRNA, complete cds.
ACCESSION U91947
VERSION U91947.1 GI:1928883
KEYWORDS
SOURCE Equus caballus (horse)
ORGANISM Equus caballus
Equus caballus
Mammalia; Eutheria; Laurasiatheria; Perissodactyla; Equidae; Equus.
REFERENCE 1 (bases 1 to 405)
AUTHORS Vandergriff,E.V. and Horohov,D.W.
TITLE Equine interleukin-5
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 405)
```



```

misc_feature 125036..138982
/note="assembly_fragment"
misc_feature 131555..172134
/note="clone overlaps with GenBank Accession Number
AC157886 clone VMRC7-34314 (center project name ilo)"
gap 138983..139082
/estimated_length=unknown
misc_feature 139083..172134
/note="assembly_fragment
clone end:other
vector_side:right"
ORIGIN
Alignment Scores:
Pred. No.: 1.13 Length: 172134
Score: 63.00 Matches: 13
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 92.86% Mismatches: 0
Query Match: 95.14% Indels: 0
DB: 14 Gaps: 0
US-10-787-382-20 (1-15) x AC158733 (1-172134)
Qy 2 AlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
Db 129022 GCTGTAGAAAGTCTCTATGATAGACTGTGTGCAGAGACCTTG 128981
RESULT 26
AC158734 177136 bp DNA linear HTG 23-MAR-2005
LOCUS Rhinolphus ferrumequinum clone VMRC7-400D16, WORKING DRAFT
DEFINITION Rhinolphus ferrumequinum (greater horseshoe bat)
ACCESSION AC158734
VERSION AC158734.1 GI:61696377
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT.
SOURCE Rhinolphus ferrumequinum (greater horseshoe bat)
ORGANISM Rhinolphus ferrumequinum
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Chiroptera; Microchiroptera;
Rhinolphidae; Rhinolphinae; Rhinolphus.
REFERENCE 1 (bases 1 to 177136)
AUTHORS Antonellis,A., Ayele,K., Benjamin,B., Blakesley,R.W., Boakye,A.,
Bouffard,G.G., Brinkley,C., Brooks,S., Chu,G., Coleman,H.,
Engle,J., Gestole,M., Greene,A., Guan,X., Gupta,J., Haghighi,P.,
Han,J., Hansen,N., Ho,S.-L., Hu,P., Hunter,G., Hurle,B., Idol,J.R.,
Kwong,P., Laric,P., Larson,S., Lee-Lin,S.-Q., Legaspi,R.,
Madden,M., Maduro,Q.L., Maduro,V.B., Margulies,E.H., Masiello,C.,
Maskeri,B., McDowell,J., Mojidi,H.A., Mullikin,J.C., Puri,O.,
Oestreicher,J.S., Park,M., Portnoy,M.E., Prasad,A., Sison,C.,
Reddix-Dugue,N., Schandler,K., Schueler,M.G., Sison,C.,
Stantripoop,S., Stephen,E., Tave,A., Thomas,J.W., Thomas,P.J.,
Tsipouri,V., Ung,L., Vogt,J.L., Wetherby,K.D., Young,A. and
Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 177136)
Green,E.D.
Direct Submission
Submitted (23-MAR-2005) NIH Intramural Sequencing Center, 5625
Fishers Lane, Rockville, MD 20852, USA
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoophgri.nih.gov
----- Project Information
Center project name: imh
Center clone name: 400D16
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 175003 bases at least Q40
Consensus quality: 175393 bases at least Q30
Consensus quality: 175604 bases at least Q20
Insert size: 219000; agarose-fp
Insert size: 176736; sum-of-contigs
Quality coverage: 8.73x in Q20 bases; agarose-fp
Quality coverage: 10.81x in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown..
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 7760: contig of 7760 bp in length
* 7761 7860: gap of unknown length
* 7861 20486: contig of 12626 bp in length
* 20487 20586: gap of unknown length
* 20587 43544: contig of 22958 bp in length
* 43545 43644: gap of unknown length
* 43645 95367: contig of 51722 bp in length
* 95367 95466: gap of unknown length
* 95467 177136: contig of 81670 bp in length.
FEATURES
source
1..177136
/organism="Rhinolphus ferrumequinum"
/mol_type="genomic DNA"
/db_xref="taxon:59479"
/clone="VMRC7-400D16"
/clone_lib="VMRC7"
/note="BAC resource: http://bacpac.chori.org/"
1..7760
/note="assembly_fragment
clone end:SP6
vector_side:right"
7761..7860
/estimated_length=unknown
7861..20486
/note="assembly_fragment"
20487..20586
/estimated_length=unknown
20587..43544
/note="assembly_fragment
clone end:other
vector_side:left"
43545..43644
/estimated_length=unknown
43645..95366
/note="assembly_fragment"
95367..95466
/estimated_length=unknown
95467..177136
/note="assembly_fragment"
ORIGIN
Alignment Scores:
Pred. No.: 1.17 Length: 177136
Score: 63.00 Matches: 13
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 92.86% Mismatches: 0
Query Match: 95.14% Indels: 0
DB: 14 Gaps: 0
US-10-787-382-20 (1-15) x AC158734 (1-177136)
Qy 2 AlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
Db 101345 GCTGTAGAAAGTCTCTATGATAGACTGTGTGCAGAGACCTTG 101386
RESULT 27
AF051372 AF051372 354 bp mRNA linear MAM 26-JUL-2000
LOCUS

```

```

DEFINITION Felis catus Interleukin-5 mRNA, partial cds.
ACCESSION AF051372
VERSION AF051372.1 GI:2961560
KEYWORDS
SOURCE
ORGANISM
Felis catus (cat)
Felis catus
Felis catus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
Felinae; Felis.
REFERENCE
1 (bases 1 to 354)
Harley,R., Helps,C.R., Harbourn,D.A., Gruffydd-Jones,T.J. and
Day,M.J.
Cytokine mRNA expression in lesions in cats with chronic
gingivostomatitis
Clin. Diagn. Lab. Immunol. 6 (4), 471-478 (1999)
10391845
REFERENCE
2 (bases 1 to 354)
Harley,R., Day,M.J., Gruffydd-Jones,T.J., Harbourn,D.A. and
Helps,C.R.
Direct Submission
Submitted (27-FEB-1998) Clinical Veterinary Science, Bristol
University, Langford House, Langford, Bristol BS40 5DU, UK
Location/Qualifiers
1..354
/organism="Felis catus"
/mol_type="mRNA"
/db_xref="taxon:9685"
/cell_line="FL4"
<1..5354
/notes="IL-5"
/codon_start=2
/product="interleukin-5"
/protein_id="AAC05752.1"
/db_xref="GI:2961561"
/translation="LGAAXVSAIAVQSPMNRVLVAETLALLSTHRTLIGDGNLMITPT
EHNHQLCIEEVFGQIDTLKNRTVPDGAVEKLFRLSLIKEHIDRQKNFGGKWKVK
NFLNLQPLGLLNT"
ORIGIN
2 AlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
61 GCTGTACAAAGTCCCATGAATAGCTGGTGGCAGACCTTG 102

Alignment Scores:
Pred. No.: 0.00383 Length: 354
Score: 60.00 Matches: 12
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 85.71% Mismatches: 0
Query Match: 81.08% Indels: 0
DB: 4 Gaps: 0

US-10-787-382-20 (1-15) x AF051372 (1-354)
Qy 2 AlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
Db 29 GCTGTACAAAGTCCCATGAATAGCTGGTGGCAGACCTTG 70

RESULT 28
AF068770
LOCUS AF068770 405 bp mRNA linear MAM 28-JUL-1998
DEFINITION Felis catus interleukin 5 mRNA, complete cds.
ACCESSION AF068770
VERSION AF068770.1 GI:3342391
KEYWORDS
SOURCE
ORGANISM
Felis catus (cat)
Felis catus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
Felinae; Felis.
REFERENCE
1 (bases 1 to 405)
Vandegrift,E., Hughes,K.J. and O'Reilly,K.L.
Direct Submission
Submitted (27-MAY-1998) VMP, LSU School of Veterinary Medicine,
South Stadium Drive, Baton Rouge, LA 70803, USA
2 (bases 1 to 405)
Vandegrift,E., Hughes,K.J. and O'Reilly,K.L.
Direct Submission

Alignment Scores:
Pred. No.: 0.00448 Length: 405
Score: 60.00 Matches: 12
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 85.71% Mismatches: 0
Query Match: 81.08% Indels: 0
DB: 4 Gaps: 0

US-10-787-382-20 (1-15) x AF068770 (1-405)
Qy 2 AlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
Db 61 GCTGTACAAAGTCCCATGAATAGCTGGTGGCAGACCTTG 102

RESULT 29
SSC010088
LOCUS SSC010088 405 bp mRNA linear MAM 15-APR-2005
DEFINITION Sus scrofa mRNA for interleukin-5.
ACCESSION AJ010088
VERSION AJ010088.2 GI:6911700
KEYWORDS IL-5 gene; interleukin-5.
SOURCE Sus scrofa (pig)
ORGANISM
Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
Sus.
REFERENCE
1
Sylvin,H., Matvienko,O., Leonchik,A., Alving,K. and van der
Ploeg,I.
Molecular cloning, expression, and purification of pig
interleukin-5
Immunogenetics 51 (1), 59-64 (2000)
10663563
2
Sylvin,H.
Direct Submission
Submitted (11-AUG-1998) Sylvin H., Department of Laboratory
Medicine, Division of Clinical Immunology, Karolinska
Institute/Karolinska Hospital, 171 76 Stockholm, SWEDEN
Revised by author
REFERENCE
3 (bases 1 to 405)
Sylvin,H.
Direct Submission
Submitted (03-FEB-2000) Sylvin H., Department of Laboratory
Medicine, Division of Clinical Immunology, Karolinska
Institute/Karolinska Hospital, 171 76 Stockholm, SWEDEN
On Feb 7, 2000 this sequence version replaced gi:6782396.
COMMENT
FEATURES
source
1..405
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
1..405
/gene="IL-5"
gene

```

```

JOURNAL Submitted (27-JUL-1998) VMP, LSU School of Veterinary Medicine,
South Stadium Drive, Baton Rouge, LA 70803, USA
REMARK Sequence update by submitter
COMMENT On Jul 28, 1998 this sequence version replaced gi:3201991.
FEATURES
Location/Qualifiers
1..405
/organism="Felis catus"
/mol_type="mRNA"
/db_xref="taxon:9685"
1..405
/codon_start=1
/product="interleukin 5"
/protein_id="AAC27616.1"
/db_xref="GI:3342392"
/translation="MRMLHLHLALGAAYVSAIAVQSPMNRVLVAETLALLSTHRTL
IGDGNLMITPTPEHNNHQLCIEEVFGQIDTLKNRTVPDGAVEKLFRLSLIKEHIDRQ
KKCGGRWRVKFLDYQLQVFLGVINTEWTIEG"
ORIGIN
US-10-787-382-20 (1-15) x AF068770 (1-405)
Qy 2 AlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
Db 61 GCTGTACAAAGTCCCATGAATAGCTGGTGGCAGACCTTG 102

Alignment Scores:
Pred. No.: 0.00448 Length: 405
Score: 60.00 Matches: 12
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 85.71% Mismatches: 0
Query Match: 81.08% Indels: 0
DB: 4 Gaps: 0

US-10-787-382-20 (1-15) x AF068770 (1-405)
Qy 2 AlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
Db 61 GCTGTACAAAGTCCCATGAATAGCTGGTGGCAGACCTTG 102

RESULT 29
SSC010088
LOCUS SSC010088 405 bp mRNA linear MAM 15-APR-2005
DEFINITION Sus scrofa mRNA for interleukin-5.
ACCESSION AJ010088
VERSION AJ010088.2 GI:6911700
KEYWORDS IL-5 gene; interleukin-5.
SOURCE Sus scrofa (pig)
ORGANISM
Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
Sus.
REFERENCE
1
Sylvin,H., Matvienko,O., Leonchik,A., Alving,K. and van der
Ploeg,I.
Molecular cloning, expression, and purification of pig
interleukin-5
Immunogenetics 51 (1), 59-64 (2000)
10663563
2
Sylvin,H.
Direct Submission
Submitted (11-AUG-1998) Sylvin H., Department of Laboratory
Medicine, Division of Clinical Immunology, Karolinska
Institute/Karolinska Hospital, 171 76 Stockholm, SWEDEN
Revised by author
REFERENCE
3 (bases 1 to 405)
Sylvin,H.
Direct Submission
Submitted (03-FEB-2000) Sylvin H., Department of Laboratory
Medicine, Division of Clinical Immunology, Karolinska
Institute/Karolinska Hospital, 171 76 Stockholm, SWEDEN
On Feb 7, 2000 this sequence version replaced gi:6782396.
COMMENT
FEATURES
source
1..405
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
1..405
/gene="IL-5"
gene

```

```

CDS
1..405
/gene="IL-5"
/codon_start=1
/product="interleukin-5"
/protein_id="CAB70611.2"
/db_xref="GI:6911701"
/db_xref="GOA:Q9MYM5"
/db_xref="InterPro:IPR000186"
/translation="MRMLHLSLGLGAAVSAIAVENTMRLVAETLTLISIHRTLL
IGDGNLMISTPVTNHQLCIEVFQGITLKNRTVPGDAVEKLFQNLSLIKEYIDRQK
KNCGERWRVTQFLDYQLQVFLGVINTWTMES"

ORIGIN
Alignment Scores:
Pred. No.: 0.00448 Length: 405
Score: 60.00 Matches: 13
Percent Similarity: 92.86% Conservative: 0
Best Local Similarity: 92.86% Mismatches: 1
Query Match: 81.08% Indels: 0
DB: 4 Gaps: 0

US-10-787-382-20 (1-15) x SSC010088 (1-405)

Qy 2 AlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
Db 61 GCTGTAGAAAATACCATGACTGGTGGCAGACCTTG 102

RESULT 30
SSC133452 Sus scrofa mRNA 529 bp mRNA linear MAM 15-APR-2005
LOCUS
DEFINITION Sus scrofa mRNA for interleukin 5.
ACCESSION AJ133452
VERSION AJ133452.2 GI:6967055
KEYWORDS IL-5 gene; interleukin 5.
SOURCE Sus scrofa (pig)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
Sus.

REFERENCE
1 Johnsen,C.K., Grondahl-Hansen,J., Johansen,R., Jungersen,G. and
Heegaard,P.M.H.
Sus scrofa mRNA for interleukin-5
Unpublished
2 Johnsen,C.K.
Direct Submission
Submitted (02-MAR-1999) Johnsen C.K., Department of Biochemistry
and Immunology, Danish Veterinary Laboratory, Bulowsvej 27, DK-1790
Copenhagen V, DENMARK
revised by [3]
3 (bases 1 to 529)
Johnsen,C.K.
Direct Submission
Submitted (07-FEB-2000) Johnsen C.K., Department of Biochemistry
and Immunology, Danish Veterinary Laboratory, Bulowsvej 27, DK-1790
Copenhagen V, DENMARK
On Feb 11, 2000 this sequence version replaced gi:4469326.

FEATURES
Location/Qualifiers
1..529
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/cell_type="lymph node"
/tissue_type="lung"
1..529
/gene="IL-5"
1..405
/gene="IL-5"
/codon_start=1
/product="interleukin 5"
/protein_id="CAB38328.1"

gene
CDS
1..405
/gene="IL-5"
/codon_start=1
/product="interleukin 5"
/protein_id="CAB38328.1"

ORIGIN
Alignment Scores:
Pred. No.: 0.0104 Length: 838
Score: 60.00 Matches: 12
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 85.71% Mismatches: 0
Query Match: 81.08% Indels: 0

/db_xref="GI:4469327"
/db_xref="GOA:Q9MYM5"
/db_xref="InterPro:IPR000186"
/db_xref="UniProt/TREMBL:Q9MYM5"
/translation="MRMLHLSLGLGAAVSAIAVENTMRLVAETLTLISIHRTLL
IGDGNLMISTPVTNHQLCIEVFQGITLKNRTVPGDAVEKLFQNLSLIKEYIDRQK
KNCGERWRVTQFLDYQLQVFLGVINTWTMES"

ORIGIN
Alignment Scores:
Pred. No.: 0.0061 Length: 529
Score: 60.00 Matches: 13
Percent Similarity: 92.86% Conservative: 0
Best Local Similarity: 92.86% Mismatches: 1
Query Match: 81.08% Indels: 0
DB: 4 Gaps: 0

US-10-787-382-20 (1-15) x SSC133452 (1-529)

Qy 2 AlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
Db 61 GCTGTAGAAAATACCATGACTGGTGGCAGACCTTG 102

RESULT 31
AF025436 Felis catus 838 bp mRNA linear MAM 20-OCT-1998
LOCUS
DEFINITION Felis catus interleukin-5 (IL-5) mRNA, complete cds.
ACCESSION AF025436
VERSION AF025436.1 GI:3228518
KEYWORDS Felis catus (cat)
SOURCE Felis catus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
Felinae; Felis.
REFERENCE
1 (bases 1 to 838)
Padrig,P.A., Qin,Y., Wells,T.N., Solway,J. and Camoretti-Mercado,B.
Sequence and structural analysis of feline interleukin-5 cDNA
Am. J. Vet. Res. 59 (10), 1263-1269 (1998)
9781459
2 (bases 1 to 838)
Padrig,P.A., Qin,Y., Wells,T.N.C., Solway,J. and
Camoretti-Mercado,B.
Direct Submission
Submitted (15-SEP-1997) Medicine, University of Chicago, 5841 S.
Maryland Avenue, Chicago, IL 60637, USA
Location/Qualifiers
1..838
/organism="Felis catus"
/mol_type="mRNA"
/db_xref="taxon:9685"
/cell_type="activated feline peripheral blood mononuclear
cells"
1..838
/gene="IL-5"
45..449
/gene="IL-5"
/note="cytokine"
/codon_start=1
/product="interleukin-5"
/protein_id="AAC64505.1"
/db_xref="GI:3228519"
/translation="MRMLHLSLGLGAAVSAIAVQSPMRLVAETLTLISIHRTLL
IGDGNLMIPTEHNNHOLCIEVFQGITLKNRTVPGDAVEKLFNLSLKEIHIDRQK
KNCGERWRVKFLDYQLQVFLGVINTWTMES"

ORIGIN
Alignment Scores:
Pred. No.: 0.0104 Length: 838
Score: 60.00 Matches: 12
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 85.71% Mismatches: 0
Query Match: 81.08% Indels: 0

```

```

DB:
US-10-787-382-20 (1-15) x AF025436 (1-838)
Qy 2 AlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
Db 105 GCTGTAGAAAGTCCCATGAATAGGCTGGTGGCAGACCTTG 146

RESULT 32
BTINTLEUS
LOCUS B.taurus mRNA for interleukin-5. 405 bp mRNA linear MAM 18-APR-2005
DEFINITION B.taurus mRNA for interleukin-5.
ACCESSION 267872
VERSION 267872.1 GI:1113120
KEYWORDS interleukin-5.
SOURCE Bos taurus (cow)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Bovinae; Bos.
REFERENCE
AUTHORS Mertens,B., Gobrecht,B. and Seow,H.F.
TITLE The nucleotide sequence of the bovine interleukin-5-encoding cDNA
JOURNAL Gene 176 (1-2), 273-274 (1996)
PUBMED 8918267
REMARK (sites)
2 (bases 1 to 405)
REFERENCE
AUTHORS Mertens,B.E.
TITLE Direct Submission
JOURNAL Submitted (13-NOV-1995) MERTENS B.E., International Livestock
Research Institute (ILRI), Bovine immunology, Naivasha road,
NAIROBI, KENYA
FEATURES
source
1..405
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/cell_type="lymphocytes"
/tissue_type="blood"
1..405
/codon_start=1
/product="interleukin-5"
/protein_id="CAA91779.1"
/db_xref="GI:1113121"
/db_xref="GOA:P52173"
/db_xref="InterPro:IPR000186"
/db_xref="InterPro:IPR009079"
/translation="MRHLTLVALGAAAYVCANAVESTMNRLLVAETLTLLSHRTL
IGDNLMIPTPQHTNHCIEEVFGQIDTLKNOTAQGDVAVKKIFQNLSLIKEIDLQK
RCGGERWRVKQFLDYLVQVPLGINTWNTMES"
ORIGIN
Alignment Scores:
Pred. No.: 0.0535 Length: 405
Score: 55.00 Matches: 12
Percent Similarity: 92.86% Conservative: 1
Best Local Similarity: 85.71% Mismatches: 1
Query Match: 74.32% Indels: 0
DB: 4 Gaps: 0

US-10-787-382-20 (1-15) x BTINTLEUS (1-405)
Qy 2 AlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
Db 61 GCTGTAGAAAGTCCCATGAATAGGCTGGTGGCAGACCTTG 102

RESULT 33
OAUI3038
LOCUS OAUI3038 520 bp mRNA linear MAM 25-MAR-1999
DEFINITION Ovis aries interleukin-5 mRNA, complete cds.
ACCESSION U35038
VERSION U35038.1 GI:4096663

DB:
US-10-787-382-20 (1-15) x OAU35038 (1-520)
Qy 2 AlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
Db 100 GCTGTAGAAAGTACCATGAATAGACTGGTGGCAGACCTTG 141

RESULT 34
OAUIV1
LOCUS Ovis aries interleukin-5 (IL5) gene, exons 1 and 2.
DEFINITION Ovis aries interleukin-5 (IL5) gene, exons 1 and 2.
ACCESSION U17052
VERSION U17052.1 GI:897558
KEYWORDS
SEGMENT 1 of 2
SOURCE Ovis aries (sheep)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Caprinae; Ovis.
REFERENCE
AUTHORS Bryson,C.E., Viney,E., Brandon,M. and Boyd,A.W.
TITLE Structure of the sheep interleukin-5 gene
JOURNAL Unpublished
2 (bases 1 to 1140)
REFERENCE
AUTHORS Bryson,C.E.
TITLE Direct Submission
JOURNAL Submitted (08-NOV-1994) Claire E. Bryson, Department of Veterinary
Science, University of Melbourne, Flemington Rd., Parkville,
Victoria, 3052, Australia
FEATURES
source
1..1140
/organism="Ovis aries"
/mol_type="genomic DNA"
/db_xref="taxon:9940"

```

CAAT_signal 339..347
/gene="IL5"
TATA_signal 389..395
/gene="IL5"
exon <470..607
/gene="IL5"
/number=1
exon 824..856
/gene="IL5"
/number=2

ORIGIN

Alignment Scores:
Pred. No.: 0.178 Length: 1140
Score: 55.00 Matches: 12
Percent Similarity: 92.86% Conservative: 1
Best Local Similarity: 85.71% Mismatches: 1
Query Match: 74.32% Indels: 0
DB: 4 Gaps: 0

US-10-787-382-20 (1-15) x OAILV1 (1-1140)

QY 2 AlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
|||||:::|
|||:::|

Dd 524 GCTGTAGAAAGTCACATGATAGACTGGTGCGAGACCTTG 565

RESULT 35
AC149665

LOCUS AC149665 linear MAM 31-JUL-2004
DEFINITION Bos taurus BAC CH240-60013 (Children's Hospital Oakland Research Institute Bovine BAC Library (male)) complete sequence.
ACCESSION AC149665
VERSION AC149665.2 GI:50872219
KEYWORDS HTG.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Bovinae; Bos.

Muzny, D., Metzker, M., Adams, C., Agbai II, O., Allen, C., Alsbrooks, S., Archer, P., Aredondo, H., Bandaranaike, D., Bangura, L., Beltran, B., Beltran, K., Beraducci, A., Biswal, K., Blyth, P., Bonham, H., Buhay, C., Burch, P., Cadoree, I., Canada, A., Cardenas, V., Carter, K., Cavazos, I., Chacko, J., Chahrouh, M., Chavez, D., Chen, A., Chen, G., Chen, R., Cheng, M.-T., Chu, J., Clerc, K., Cockrell, R., Coyle, M., Cree, A., Curry, S., Dai, W., Davila, M.L., Davis, C., Davy-Carroll, L., De And, C., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Donlin, J., McCauley, S., Dugan-Rocha, S., Dunn, A., Durbin, K., Dziuda, D., Egan, A., Escotto, M., Espinosa, V., Eugene, C., Fa, M., Fernandez, S., Fernando, P., Flagg, N., Forbes, L., Foster, P., Fowler, G., Fu, Q., Fuh, E., Garcia, A., Garcia, R., Garner, T., Gaslin, C., Gench, S., Ghose, S., Gill, R., Gonzalez, D., Gonzalez-Garay, M., Guevara, W., Holder, R., Haaland, W., Haebleren, K., Hall, B., Hamid, H., Hamilton, K., Harbes, B., Harris, R., Havila, P., Hawes, A., Hawkins, E., Hayes, S., Hemphill, L., Hernandez, J., Hines, S., Hitchens, M., Hodgson, A., Hogues, M., Hollins, B., Howell, L.T., Huliyk, S., Hume, J., Imo, K., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Kalafus, K., Kelly, S., Keys, T., Khan, Z., King, D., Kovar, C., Kowis, A., Kowis, C., Lara, F., Leal, S., Lee, K., Lee, S., LeCall, F.I., Lemon, S., Lewis, L., Li, B., Li, Y., Li, Z., Linnell, M., Liu, W., Liu, Y.-S., Liu, Y., Liyanage, D., London, P., Lopez, J., Lorenshewala, L., Lozado, R., Luk, T., Madu, R., Maheshwari, M., Mahoney, C., Malloy, K., Mansouri, D., Martinez, E., McCelland, H., McPherson, J., Mercadado, C., Milosavljevic, A., Minja, E., Morgan, M., Morris, S., Munidasa, M., Murray, D., Nazareth, L., Ngo, D., Nguyen, N., Norwig-Eastaugh, E., Nott, A., Nwaokemele, O., Obregon, M., Ochi-Okorie, C., Odeh, E., Okwuonu, G., Okwuonu, K., Parker, D., Pasternak, S., Patel, B., Patel, V., Paul, H., Perez, A., Perez, L., Petrosino, J., Pham, T., Primus, E., Pu, L.-L., Puafo, M., Qin, X., Quinn, A., Quiroz, J., Rabata, D., Rachlin, E., Reigh, R., Ren, Y., Reuter, M., Richards, S., Rives, C., Rodriguez, F., Rojas, A., Ruiz, S.J., Sana, M., Sanders, W., Santibanez, J., Santos, R.,

repeat_region /rpt family="Bov-tA2" 11219.11498
repeat_region /rpt family="BOV-A2" 11588.11178
repeat_region /rpt family="L1MA10" 12948.12981
repeat_region /rpt family="(TTTA)n" complement(12982.13213)
repeat_region /rpt family="CHR-2B" 13262.13466
repeat_region /rpt family="Bov-tA2" 13521.13734
repeat_region /rpt family="Bov-tA1" 14293.14578
repeat_region /rpt family="CHR-2A" 14718.14815
repeat_region /rpt family="L1MC/D" 14830.15030
repeat_region /rpt family="Bov-tA1" 15418.15448
repeat_region /rpt family="(A)n" 15508.15732
repeat_region /rpt family="Bov-tA2" complement(16078.16128)
repeat_region /rpt family="Bov-tA2" complement(16359.17032)
repeat_region /rpt family="L1MC3" complement(17029.17089)
repeat_region /rpt family="L1_Art" 17205.17298
repeat_region /rpt family="L1_Art" 17380.17404
repeat_region /rpt family="AT_rich" 17574.18113
repeat_region /rpt family="L1MB3" 18229.18475
repeat_region /rpt family="MER102b" 19565.19653
repeat_region /rpt family="MLT1K" complement(19660.19862)
repeat_region /rpt family="Bov-tA3" complement(21061.21617)
repeat_region /rpt family="L1MA5" 21810.21836
repeat_region /rpt family="AT_rich" 21841.22126
repeat_region /rpt family="BOV-A2" complement(22186.22227)
repeat_region /rpt family="BOV-A2" 23038.23066
repeat_region /rpt family="(TA)n" 23107.23134
repeat_region /rpt family="AT_rich" 23458.23485
repeat_region /rpt family="(TGAA)n" 23927.24147
repeat_region /rpt family="CHR-2A" complement(24678.24775)
repeat_region /rpt family="L2" complement(24777.24979)
repeat_region /rpt family="Bov-tA1" complement(24911.24996)
repeat_region /rpt family="CHR-2A" 24997.25045
repeat_region /rpt family="(TA)n" complement(25052.25309)
repeat_region /rpt family="BOV-A2" 25335.25673
repeat_region /rpt family="L1MA9" 25769.25797
repeat_region /rpt family="AT_rich" 26488.27365
repeat_region /rpt family="Bov-B"

repeat_region complement(28556.28630)
repeat_region /rpt family="tRNA-Glu-GAG" complement(28698.28784)
repeat_region /rpt family="MLT11" 29385.30014
repeat_region /rpt family="L1_Art" 30031.30105
repeat_region /rpt family="MER91B" 30141.30298

Alignment Scores:
Pred. No.: 70.3 Length: 197131
Score: 55.00 Matches: 12
Percent Similarity: 92.86% Conservative: 1
Best Local Similarity: 85.71% Mismatches: 1
Query Match: 74.32% Indels: 0
DB: 4 Gaps: 0

US-10-787-382-20 (1-15) x AC149665 (1-197131)

QY 2 AlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
Db 20446 GCTGTAGAAAGTACCATGATAGACTGGTGCAGAGACCTTG 20487

RESULT 36
CR543861.14/c
WPCOMMENT

Sequence split into 36 fragments LOCUS CR543861 Accession CR543861

Fragment Name	Begin	End
CR543861_00	1	110000
CR543861_01	100001	210000
CR543861_02	200001	310000
CR543861_03	300001	410000
CR543861_04	400001	510000
CR543861_05	500001	610000
CR543861_06	600001	710000
CR543861_07	700001	810000
CR543861_08	800001	910000
CR543861_09	900001	1010000
CR543861_10	1000001	1110000
CR543861_11	1100001	1210000
CR543861_12	1200001	1310000
CR543861_13	1300001	1410000
CR543861_14	1400001	1510000
CR543861_15	1500001	1610000
CR543861_16	1600001	1710000
CR543861_17	1700001	1810000
CR543861_18	1800001	1910000
CR543861_19	1900001	2010000
CR543861_20	2000001	2110000
CR543861_21	2100001	2210000
CR543861_22	2200001	2310000
CR543861_23	2300001	2410000
CR543861_24	2400001	2510000
CR543861_25	2500001	2610000
CR543861_26	2600001	2710000
CR543861_27	2700001	2810000
CR543861_28	2800001	2910000
CR543861_29	2900001	3010000
CR543861_30	3000001	3110000
CR543861_31	3100001	3210000
CR543861_32	3200001	3310000
CR543861_33	3300001	3410000
CR543861_34	3400001	3510000
CR543861_35	3500001	3598621

Continuation (15 of 36) of CR543861 from base 1400001 (CR543861 Acinetobacter sp. ADP1)

Alignment Scores:
Pred. No.: 702 Length: 110000
Score: 49.00 Matches: 10
Percent Similarity: 86.67% Conservative: 3
Best Local Similarity: 66.67% Mismatches: 2
Query Match: 66.22% Indels: 0


```

DB: 1 Gaps: 0
US-10-787-382-20 (1-15) x CR543861_14 (1-110000)
QY 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
    |||||::: ||:::|||||:::|||||
Db 73554 TTTCAGTGAAGGTACCACTTAACAGATTGCTTGACAGCCGTTA 73510

RESULT 37
AF419552/c 1247 bp mRNA linear PLN 15-NOV-2001
LOCUS Arabidopsis thaliana At2g47020/F14M4.15 mRNA sequence.
ACCESSION AF419552
VERSION AF419552.1 GI:16930397
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1 (bases 1 to 1247)
AUTHORS Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J.,
Bower,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G.,
Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M.,
Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K.,
Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C.,
Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K.,
Davis,R.W., Theologis,A. and Ecker,J.R.
TITLE Arabidopsis cDNA clones
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1247)
AUTHORS Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J.,
Bower,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G.,
Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M.,
Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K.,
Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C.,
Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K.,
Davis,R.W., Theologis,A. and Ecker,J.R.
TITLE Direct Submission
JOURNAL Submitted (14-SEP-2001) Salk Institute Genomic Analysis Laboratory
(SIGNAL), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA
COMMENT Incomplete ORF.
RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.
The Salk, Stanford, PGEN (SSP) Consortium members carried out the
sequencing and annotation of the RAFL cDNAs: Cheuk,R., Chen,H.,
Kim,C.J., Koesema,E., Meyers,M.C., Shinn,P., Banh,J., Bower,L.,
Dale,J.M., Goldsmith,A.D., Jiang,P.X., Jones,T., Karlin-Neumann,G.,
Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Nguyen,M.,
Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Southwick,A.,
Tang,C.C., Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S.,
Davis,R.W., Theologis,A. and Ecker,J.R.
Cheuk,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to
this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)
contributed equally to this work as PIs.
FEATURES
Location/Qualifiers
1..1247
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="2"
/ecotype="Columbia"
source
ORIGIN

```

```

Alignment Scores:
Pred. No.: 10.5 Length: 1247
Scores: 47.00 Matches: 9
Percent Similarity: 90.00% Conservatives: 0
Best Local Similarity: 90.00% Mismatches: 1
Query Match: 63.51% Indels: 0
DB: 15 Gaps: 0

US-10-787-382-20 (1-15) x AF419552 (1-1247)
QY 5 AsnProMetAsnArgLeuValAlaGluThr 14
    |||||::: ||:::|||||:::|||||
Db 519 AATCCCATGAACAGATTGGCAGCGAGACC 490

RESULT 38
AC004411 106329 bp DNA linear PLN 11-MAR-2002
LOCUS Arabidopsis thaliana chromosome 2 clone F14M4 map CIC06C03,
DEFINITION complete sequence.
ACCESSION AC004411
VERSION AC004411.3 GI:20197133
KEYWORDS HTG.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1 (bases 1 to 106329)
AUTHORS Rounsley,S.D., Lin,X., Kaul,S., Shea,T.P., Fujii,C.Y., Mason,T.M.,
Shen,M., Romning,C.M., Fraser,C.M., Somerville,C.R. and Venter,J.C.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 106329)
AUTHORS Lin,X.
TITLE Direct Submission
JOURNAL Submitted (09-MAR-2000) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA
REFERENCE 3 (bases 1 to 106329)
AUTHORS Town,C.D. and Kaul,S.
TITLE Direct Submission
JOURNAL Submitted (27-FEB-2002) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA, cdtown@igr.org
COMMENT On Apr 18, 2002 this sequence version replaced gi:6598415.
FEATURES
Location/Qualifiers
1..106329
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="2"
/map="CIC06C03"
/clone="F14M4"
/ecotype="Columbia"
misc_feature
1..4507
/note="overlap with BAC clone T3D7
(AC007236:11525..16031)."
repeat_region
complement(3345..3372)
/rpt_family="AT-rich"
repeat_region
complement(4880..4962)
/rpt_family="AT-rich"
repeat_region
complement(6162..6182)
/rpt_family="(TAA)n"
6464..7254
/gene="AT2g47150"
/note="synonym: F14M4.2"
mRNA
join(<6464..6654,6740..6964,7068..7254)
/gene="AT2g47150"
CDS
join(6464..6654,6740..6964,7068..7254)
/gene="AT2g47150"
/codon_start=1
/product="putative alcohol dehydrogenase"
/protein_id="AAC34216.1"
/db_xref="GI:3522934"
/translation="MDGKIVITGGASGTGAESARLFTDHGAQVVVVVQLQBEQGTSP

```

gene
FOSAKTEQVTVMLQTRRNOPGVLETRGSIIDLNLRLERFHRMTAVNVRGAASIKHAA
RAMVKGRTGSIVCTSTSEIVVRDLNTRRRSMGSHDEETAKEATCEYCEARGIFPKG
VVKARHVAEAAFLASDDSVIISQNLNLAVDGGFCVVKPI"
7490. .8838
/gene="At2g47140"
/note="synonym: F14M4.3; supported by cDNA:
gi_13878112.gb.AF370319.1.AF370319"
Join(7490. .7532,7626. .8078,8302. .8838)
/gene="At2g47140"
Join(7519. .7532,7626. .8078,8302. .8608)
/gene="At2g47140"
/codon_start=1
/product="putative alcohol dehydrogenase"
/protein_id="AAC34234.2"
/db_xref="GI:20197135"
/translation="MSGKRLDGKIVIIITGGASGIGAESVRLTFEHRGARVIVDVQDEL
GNVAVSGEDKASYHCDVTNTEVENAVKFTVBYKGLDVLFSNAGVIEPFSIL
LNLNLDRTIAINLRTAAFIKHAARAVKEGIRGIVCTTSVAAEIAGTAPHGVTTS
KHLLGLIKASGGLGKYGIRVNGVAPFGVATPLVCGPFKMEPNVVEQNTSASANLKG
TVLKARHVAEAAFLASDDSAVVSQNLNLAVDGGYSVVKP"
9045. .9951
/gene="At2g47130"
/note="synonym: F14M4.4"
Join(9045. .9110,9192. .>9951)
/gene="At2g47130"
Join(9097. .9110,9192. .9951)
/gene="At2g47130"
/codon_start=1
/product="putative alcohol dehydrogenase"
/protein_id="AAC34217.1"
/db_xref="GI:3522935"
/translation="MSGRLRDLGKIALIITGGASGIGAEAVRLTFDHGAKVIVDVQBEL
GNVAVSGKDKASFRCDVTNEKEVENAVKFTVBYKGLDVLFSNAGVIEPFSFLD
LNLNLDRTIAINLRTAAFIKHAARAVKEGIRGIVCTTSVAEIEGGPHAYTAS
KHALGLIKASGGLGKYGIRVNGVAPFGVATPLVCGPFKMEPNVVEQNTSAAATGILKV
VVKARHVAEAAFLASDDSAVVSQNLNLAVDGGYSVVKPI"
11340. .12219
/gene="At2g47120"
/note="synonym: F14M4.5"
Join(11340. .11368,11457. .>12219)
/gene="At2g47120"
Join(11355. .11368,11457. .12219)
/gene="At2g47120"
/codon_start=1
/product="putative alcohol dehydrogenase"
/protein_id="AAC34218.1"
/db_xref="GI:3522936"
/translation="MSGRLRDLGKIVIIITGGASGIGADAAARLTFDRGAKVIVDVQBEL
GNVAVLIGDKKASFRCDVTNTEVEDAVKFTVEKHGKLDVLFSNAGVIEPFSFLD
FLERFDRITAVNVRGAAFIKHAARAVKEGIRGIVCTTSVAEIEGGHGYTASK
HGLVGLIRACSDGLGKYGIRVNGVAPFGVATPLVCGPFKMEPNVVEQNTSAAATGILKV
LKASHVAQVALFLASDDSAVVSQNLNLAVDGGYTVVKPSRD"
12714. .14202
/gene="At2g47115"
/note="synonym: F14M4.47"
Join(<12714. .13024,13117. .13150,13223. .13339,13458. .13547,
13627. .13692,13770. .13861,13947. .13994,14091. .>14202)
/gene="At2g47115"
Join(12714. .13024,13117. .13150,13223. .13339,13458. .13547,
13627. .13692,13770. .13861,13947. .13994,14091. .14202)
/gene="At2g47115"
/codon_start=1
/product="predicted protein"
/protein_id="AAW14934.1"
/db_xref="GI:20197140"
/translation="NALIVYWDYDFCFAIVAAIVTSLWFLSRRDRGCVVIDTDSHDS
LPLPLVRNRGARGLWASQWTRLHPGWLFTSTSPLSMAALLAWDVIKWDASIFVY
TWTFMNLVIFYFMGIVASVYGLIHLKELTLETDEDVIVKVGDEFRRLTSAGV
VLTDIVFWLIVVFFSTTRFGNTLTIOMHTANAGLLETLNLSLPFFWFMGIFVL
WSCLYVIFQWIIHACGFTWPPYFLEDKPAPIWYLCMAIWHIIPCYGAYAAIVKAKN
complement(14282. .14910)
/gene="At2g47110"

gene
/note="synonym: F14M4.6; identical to GI:166936; supported
by full length cDNA: Ceres:6247"
complement(<14282. .>14910)
/gene="At2g47110"
complement(14404. .14877)
/gene="At2g47110"
/codon_start=1
/product="ubiquitin extension protein (UBQ6)"
/protein_id="AAC34235.1"
/db_xref="GI:3522953"
/translation="MQIFVKLTGTKTIILEVSSDTIDNVAKIODEKGIIPDQORLI
PAGQLEDGRTLADYNIQKSTLHLVLRGCAKKRKKTYTKPKIKHKHKVLA
LQYKVDGSGKVLRLKECPNATCGAGTFMASHFDRHYCGKGLTYVYQKEGAE"
complement(15079. .15151)
/gene="At2g47100"
complement(15079. .15151)
/gene="At2g47100"
/product="crNA-Ala"
15281. .18297
/gene="At2g47090"
/note="synonym: F14M4.8"
Join(<15281. .15448,15590. .15898,15984. .16192,16274. .16322,
16393. .16482,16722. .18297)
/gene="At2g47090"
Join(15281. .15448,15590. .15898,15984. .16192,16274. .16322,
16393. .16482,16722. .18134)
/gene="At2g47090"
/codon_start=1
/product="unknown protein"
/protein_id="AAC34219.2"
/db_xref="GI:20197141"
/translation="MDSCAVCAENLWVGSGCHREVCTCVVRLFIILNDRCCI
KCTCEPVVFTKALGDYTKTISDSTTFPPVPKGRVGSFWYHETNVPFDLNYTR
IKAMRCSLNCNDNTNPKPKKPNCHVFRKSVHELKDLHAKHLMCSLCLGVKKV
FICEQKLTQGLNQHISSGSEVDGSESGEFTGHPMCEFCRPFYGDNELYTHMS
REHYTCHICQLKPGQYEGYNDLEVFHRSDFHFLCEDETCLAKKFTVFOIEAELKI
QASFOYPNRRRRSLRLENLVLLESOASYAENDNLLPOHVRSGNSRLGESFP
PLVQAQOQSRFGQNSLVSNTTTFQHRANQOQSRFGQNSLVSNTTTFQRH
QTNRSATSGSQAMPALNRGPABEISITSRVSSGASQSQSRHRDRVSTRLASAVP
QDARTTGGSSGSLSSANATKRNHHSSSTPKMSETRSLAQPHSDSPQISAVNR
RSSSTANAGNIQVAGVSDQSNKSLVERIKHASLGHDEELFWAFKNTSGKYRHSI
DARTYLEYKGYGLSHLVDMLARCPDQROKELIDITHNACLKGNKGAKVBESSSD
SKGDRFVTVRKLOPDSKQDKDKDAYRDKGKTVTTLVNSSAGVGLDTCQKP
KTKSKFLTRLGKSMAAVLDLRNSNPEPEPKNDNSKRSQNSPGLPLRGAWKRG
AKLRY"
complement(18862. .20346)
/gene="At2g47080"
/note="synonym: F14M4.9"
complement(join(<18862. .19451,19533. .20078,20167. .>20346))
18867. .18897
/rpt_family="CAA)n"
complement(join(19026. .19451,19533. .20078,20167. .20346))
/gene="At2g47080"
/codon_start=1
/product="unknown protein"
/protein_id="AAC34220.1"
/db_xref="GI:3522938"
/translation="MSLYFLRYLLCSVEGKYLIIQETHDSTTREDDDKFONSEIVECV
NFCSDMPILSGRGMIEIDQGLSSFFFLVVEDDDVCEIRILETTLEFTGDSAKQ
AMDFTHEIGLLHRSKLGESDPNPGVPLIRFQWLIETFSMDREWCATVIRKLNMFPDG
AVGEFSSSNATLSLCELHRAVRKNSKPMVEMLLARYPKOORNSLFRPDAAGPAGLT
PLHIAAGDGSDELVDLTEDPAMVGIENAKTCRSTGTFTEDYARLGHGFHYIHLIQ
RKIAKSTTDEHVVNIIPVPSDFRQKPKPGPMASAEITQIPCKLCKHLVYIGTTR
RSVAYPAMLNSMAIAAVCVVALLFKSCPEVLYVFPQFRELIDYGS"
complement(21637. .22434)
/gene="At2g47070"
/note="synonym: F14M4.10"
complement(join(<21637. .21904,21980. .>22434))
/gene="At2g47070"
complement(join(21637. .21904,21980. .>22434))

Alignment Scores:

Pred. No.: 1.82e+03 Length: 106329
 Score: 47.00 Matches: 9
 Percent Similarity: 90.00% Conservative: 0
 Best Local Similarity: 90.00% Mismatches: 1
 Query Match: 63.51% Indels: 0
 DB: 15 Gaps: 0

US-10-787-382-20 (1-15) x AC004411 (1-106329)

Qy 5 AsnProMetAsnArgLeuValAlaGluThr 14

Db 39280 AATCCCATGACAGATTGCACGGAGACC 39309

RESULT 39
 SCO939129/c

LOCUS SCO939129 292200 bp DNA linear BCT 16-APR-2005
 DEFINITION Streptomyces coelicolor A3(2) complete genome; segment 26/29.
 ACCESSION AL049485 AL049487 AL079355 AL138598 AL138667 AL138668
 AL138977 AL138978 AL138979 AL138980 AL138981 AL138982
 AL939129.1 GI:24418971

VERSION Streptomyces coelicolor A3(2)

KEYWORDS Streptomyces coelicolor A3(2)

SOURCE Streptomyces coelicolor A3(2)

ORGANISM Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

REFERENCE Streptomycineae; Streptomycetaceae; Streptomyces.

1 Bentley, S.D., Chater, K.F., Cerdeno-Tarraga, A.M., Challis, G.L.,

Thompson, N.R., James, K.D., Harris, D.E., Quail, M.A., Kieser, H.,

Harper, D., Bateman, A., Brown, S., Chandra, G., Chen, C.W., Collins, M.,

Cronin, A., Fraser, A., Goble, A., Hidalgo, J., Hornsby, T., Howarth, S.,

Huang, C.H., Kieser, T., Larke, L., Murphy, L., Oliver, K., O'Neill, S.,

Rabinowitch, E., Rajandream, M.A., Rutherford, K., Rutter, S.,

Seeger, K., Saunders, D., Sharp, S., Squares, R., Squares, S.,

Taylor, K., Warren, T., Wietzorrek, A., Woodward, J., Barrell, B.G.,

Parkhill, J. and Hopwood, D.A.

Complete genome sequence of the model actinomycete Streptomyces

coelicolor A3(2)

Nature 417 (6885), 141-147 (2002)

12000953

2 (bases 1 to 292200)

Bentley, S.D.

Direct Submission

Submitted (09-MAY-2002) Submitted on behalf of the Streptomyces

sequencing team, Sanger Institute, Wellcome Trust Genome Campus,

Hinxton, Cambridge CB10 1SA E-mail: sdb@sanger.ac.uk

On or before Oct 29, 2002 this sequence version replaced

gi:20520762, gi:20520792, gi:20520672, gi:20520776, gi:20520777,

gi:20520778, gi:20520779, gi:20520673, gi:20520674.

Location/Qualifiers

1..292200

/organism="Streptomyces coelicolor A3(2)"

/mol_type="genomic DNA"

/strain="A3(2)"

/db_xref="taxon:100226"

58..1089

/gene="SCO6712"

/note="synonym: SC4C6.22"

58..1089

/gene="SCO6712"

/note="SC4C6.22, possible copper oxidase, len: 343 aa;

weakly similar to members of the multicopper oxidase

family e.g. to the N-terminus of SW:COFA_PSES8

(EMBL:MI19930), CopA, Pseudomonas syringae copper

resistance protein A precursor (609 aa), fasta scores:

opt: 170 z-score: 188.4 E(): 0.0036, 27.3% identity in 205

aa overlap. Also weakly similar to the blood coagulation

factors which are structurally related to the multicopper

oxidases. Similar to the N-terminus of TR:O53858

(EMBL:AL022004) Mycobacterium tuberculosis hypothetical

protein (504 aa) (25.1% identity in 323 aa overlap).

Contains PS00079 Multicopper oxidases signature 1 and

PS00080 Multicopper oxidases signature 2 (the latter

suggesting copper-binding)."

/codon_start=1
 /transl_table=11
 /product="putative copper oxidase"
 /protein_id="CAB45586.1"
 /db_xref="GI:5139568"
 /db_xref="GOA:Q9XAL8"
 /db_xref="InterPro:IPR002335"
 /db_xref="InterPro:IPR006311"
 /db_xref="InterPro:IPR011706"
 /db_xref="UniProt/TREMBL:Q9XAL8"
 /translation="MDRRGGRNRVLLGGAATSLSTAPEVAGAAPAAKGIATARTAPA
 GGEVRHLKMYAEKLADGOMGVGFEKGRASVPGPLIEVNEGDTLHLEFNTMDVRASLH
 VHGLDISSDGTAMKSDVEPGGTRTYTTRTHKGRDDGTWRPGSAGYHYHHDVV
 GTEHGTGIRNGLYGPVIVRKGDLVDPATHITVFNMDTINNRKHTGPDDEATVGDR
 VEIYMTGHEYYTHFMHGRWADNRGTILTGPDPPDPRVIDNKTIGPADSGFQIIAG
 EGVGAWMTYHCHVQSHSDMGVGLFLVKPDGTIPGYEPHEHGATAKSGESGPTG
 GAAAEHEH"
 521..544
 /gene="SCO6712"
 /note="repeat 7; 23 bp imperfect inverted repeat"
 901..963
 /gene="SCO6712"
 /note="PS00079 Multicopper oxidases signature 1"
 916..951
 /gene="SCO6712"
 /note="PS00080 Multicopper oxidases signature 2"
 1276..2334
 /gene="SCO6713"
 /note="synonym: SC4C6.23"
 1276..2334
 /gene="SCO6713"
 /note="SC4C6.23, probable lacI-family transcriptional
 regulator, len: 352 aa; similar to many e.g. TR:O87590
 (EMBL:AF086819), CelR, Thermomonospora fusca
 transcriptional regulator of cellulase genes (340 aa),
 fasta scores; opt: 991 z-score: 1119.2 E(): 0, 51.9%,
 identity in 324 aa overlap. Similar to others from S.
 coelicolor e.g. TR:O86795 (EMBL:AL031317) S. coelicolor
 putative transcriptional regulator (355 aa) (36.3%
 identity in 342 aa overlap). Contains Pfam matches to
 entry PF00532 Peripla BP like, Periplasmic binding
 proteins and LacI family and to entry PF00356 lacI,
 Bacterial regulatory proteins, lacI family. Contains
 probable helix-turn-helix motif at aa 9-30 (Score 2307,
 +7.04 SD)"
 /codon_start=1
 /transl_table=11
 /product="putative transcriptional regulator"
 /protein_id="CAB45587.1"
 /db_xref="GI:5139569"
 /db_xref="GOA:Q9XAL7"
 /db_xref="InterPro:IPR001761"
 /db_xref="InterPro:IPR000843"
 /db_xref="UniProt/TREMBL:Q9XAL7"
 /translation="MTETGPRPTLEVAERAGVSRATVSRVVGNGHGVDRDALEARRVR
 AVEELGYVNHAAEVLTRRHOAVVAAPETRVFADPYFAQLRGISKELTAHDNQ
 LVLLITGREDHARVGRYLAGGHVGLVSLHLHDLPLGLVRSAGVPTVFGGRPDWD
 DGRDDVYVDSDNRGGARSVHVLGALGRTHIAHTGPDQTSADRAGRDVRPDGA
 GPGIVARGDFTSGGGERARELLDRCPGLDAVFAANDLTAAGALRVLRERGRVDPDV
 AVWGFDMLPVAEQTPDPTTVTRQDIEGMGLMARLLLRGLDRRAADATDFAARTPT
 AAPFGVLLPTTLVHRSTA"
 1291..1374
 /gene="SCO6713"
 /note="Pfam match to entry PF00356 lacI, Bacterial
 regulatory proteins, lacI family, score 39.70, E-value
 9.4e-10"
 1465..2202
 /gene="SCO6713"
 /note="Pfam match to entry PF00532 Peripla BP like,
 Periplasmic binding proteins and LacI family, score
 95.60, E-value 9.5e-25"
 complement(2088..2111)
 /note="repeat 7; 23 bp imperfect inverted repeat"

```

gene
complement(2440. .3231)
/ gene="SCO6714"
/ note="synonym: SC4C6.24c"
CDS
complement(2440. .3231)
/ gene="SCO6714"
/ note="SC4C6.24c, possible hydroxylase, len: 263 aa;
similar to TR:Q55078 (EMBL:U50973) Streptomyces sp. strain
C5 daunomycin C-14 hydroxylase (275 aa), fasta scores;
opt: 422 z-score: 482.3 E0 : 1.5e-19, 38.5% identity in
273 aa overlap. Similar to TR:O50527 (EMBL:AL009204)
S.coelicolor hypothetical protein (267 aa) (37.9% identity
in 253 aa overlap) and to (EMBL:AL049863), SCSh1.09c.
S.coelicolor possible hydroxylase (265 aa) (35.7% identity
in 258 aa overlap). Also similar to hypothetical proteins
from Mycobacterium tuberculosis and to Mycobacterium
tuberculosis 27.3 kDa WAB HBT7 reactive antigen
(EMBL:AJ007737) (260 aa) (36.7% identity in 264 aa
overlap)"
/ codon_start=1
/ transl_table=11
/ product="putative hydroxylase"
/ protein_id="CAB45588.1"
/ db_xref="GI:5139570"
/ db_xref="InterPro:IPR004360"
/ db_xref="UniProt/TrEMBL:Q9XAL6"
/ translation="MLTTRFVTVGAPNLVDVGTSLDGLGATSYGGLFGWRFSRAGPEAG
GVAFELDGRITVAGQGGTTEBQPPSTVTVFOAPDARAQAAGAEHGGVLVQPMQDV
DOGTMALADRAGLPGIWMQPGQAGLDVDTGSGALCWELHTADTAAAAAYVRAVLG
LETSGVSFPGSGYTCVPAGSEGDAMFGGLVPLAEDPADTDADACAGLFPFVAVDDADTA
VARTTELGSTVRMPATIDIEGVGRVARLADPYGARFVLRPAAPRQG"
RBS
complement(3239. .3244)
/ note="possible RBS"
/ gene="SCO6714"
/ note="synonym: SC4C6.25"
/ note="synonym: SC4C6.25"
/ note="synonym: SC4C6.25"
/ gene="SCO6715"
/ note="SC4C6.25, probable transcriptional regulator, len:
81 aa, similar to many transcriptional regulators (some
putative) from Streptomyces e.g. TR:O53963 (EMBL:X62287),
WhbB, S.coelicolor transcriptional regulator essential for
sporulation (87 aa), fasta scores; opt: 336 z-score: 452.7
E0 : 5.7e-18, 61.1% identity in 72 aa overlap. Also
similar to e.g. TR:O59649 (EMBL:AL02212) Mycobacterium
tuberculosis putative regulatory protein (100 aa) (40.0%
identity in 75 aa overlap)"
/ codon_start=1
/ transl_table=11
/ product="putative transcriptional regulator"
/ protein_id="CAB45589.1"
/ db_xref="GI:5139571"
/ db_xref="GOA:Q9XAL5"
/ db_xref="InterPro:IPR003482"
/ db_xref="UniProt/TrEMBL:Q9XAL5"
/ translation="WHTETTTAPDLSMOEALCAQTGGDFFFPFPGSSVRDAKRICRAL
CPIRSTCLEFALLSNDERFGVGGGLSEKRLALRRTP"
gene
4003. .4197
/ gene="SCO6716"
/ note="SC4C6.26, hypothetical protein, len: 64 aa; unknown
function, probable CDS suggested by positional base
preference, GC frame plot and amino acid composition"
/ codon_start=1
/ transl_table=11
/ product="hypothetical protein"
/ protein_id="CAB45590.1"
/ db_xref="GI:5139572"

```

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
gene
CDS
ORIGIN
Alignment Scores:
Pred. No.:
Score:
Percent Similarity:
Best Local Similarity:
Query Match:
DB:
US-10-787-382-20 (1-15) x AF148211 (1-448)
Qy 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
Db 74 TTTCGCTGTGGAGATCCCTATGACACCGGTAGTGAAGAGACCTTG 118
RESULT 42
AX250264
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. .13600
/organism="Streptomyces cinnamonensis"
/mol_type="unassigned DNA"
/db_xref="taxon:1900"

Sigmodon hispidus (hispid cotton rat)
Sigmodon hispidus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Cricetidae; Sigmodontinae; Sigmodon.
1 (bases 1 to 448)
Howard, S., Jacquet, A., Haumont, M., Daminet, V., Milican, F.,
Glineur, F. and Bollen, A.
Cloning, expression and purification of recombinant cotton rat
interleukin-5
Gene 257 (1), 149-155 (2000)
11054577
2 (bases 1 to 448)
Howard, S.
Direct Submission
Submitted (03-MAY-1999) Applied Genetics, Free University of
Brussels, rue de l'Industrie, 24, Nivelles 1400, Belgium
Location/Qualifiers
1. .448
/organism="Sigmodon hispidus"
/mol_type="mRNA"
/db_xref="taxon:42415"
/tissue_type="spleen"
1. .448
/gene="IL-5"
23. .421
/gene="IL-5"
/note="cytokine"
/codon_start=1
/product="interleukin-5"
/protein_id="AAG16722.1"
/db_xref="GI:10334680"
/translation="MRMLHLISILFLACWTFVAVEIPMHTVVTVKETLIQLSTHALLTS
NETVLPTVTHKHOLCIETIFRGIDILKNQTVRGTVETILFQNLILKKYIDRQKEK
CGEERRRFRQLDYLOEFLVGMGTWTEH"

5.24 Length: 448
46.00 Matches: 10
80.00% Conservative: 2
66.67% Mismatches: 3
62.16% Indels: 0
9 Gaps: 0

AX250264 13600 bp DNA linear PAT 05-OCT-2001
Sequence 4 from Patent WO0168867.
AX250264
AX250264.1 GI:15984066
Streptomyces cinnamonensis
Streptomyces cinnamonensis
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Streptomycineae; Streptomycetaceae; Streptomycetes.
1
Leadlay, P.P., Staunton, J. and O'Liynyk, M.C.
Polyketides and their synthesis
Patent: WO 0168867-A 4 20-SEP-2001;
Biotica Technology Limited (GB)
Location/Qualifiers
1. .13600
/organism="Streptomyces cinnamonensis"
/mol_type="unassigned DNA"
/db_xref="taxon:1900"

/note="Nucleotides 90001 - 103600 of the monensin
biosynthetic gene cluster"
ORIGIN
Alignment Scores:
Pred. No.:
Score:
Percent Similarity:
Best Local Similarity:
Query Match:
DB:
US-10-787-382-20 (1-15) x AX250264 (1-13600)
Qy 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGlu 13
Db 8150-TTCTCTGTGCACGAGCGATGTCACGTCTGTGCGCGAA 8188
RESULT 43
AC014950
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
1. .16535
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"

AC014950 16535 bp DNA linear HTG 16-NOV-1999
Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***
AC014950
AC014950.1 GI:6436385
HTG; HTGS_PHASE2.
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 16535)
Adams, M. and Venter, J.C.
Direct Submission
Submitted (16-NOV-1999) Celera Genomics, 45 West Gude Drive,
Rockville, MD, USA
This sequence was identified as CDM:10210325 by the submitter.
For further information on this sequence e-mail to fly@celera.com.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
Location/Qualifiers
1. .16535
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"

Alignment Scores:
Pred. No.:
Score:
Percent Similarity:
Best Local Similarity:
Query Match:
DB:
US-10-787-382-20 (1-15) x AC014950 (1-16535)
Qy 1 PheAlaValGluAsnProMetAsnArgLeuValAla 12
Db 15098 TTTCGCTGTGCAAAACCCATTAAACCGATTGATCGG 15133
RESULT 44
AB071139
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. 17525 bp RNA linear INV 21-MAY-2003

AB071139 17525 bp RNA linear INV 21-MAY-2003
Apis mellifera Ks-1 non-coding nuclear RNA.
DEFINITION
AB071139
AB071139.1 GI:22138802
Apis mellifera (honey bee)
Apis mellifera
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
Neoptera; Endopterygota; Hymenoptera; Apoidea; Apoidea;

```

REFERENCE
AUTHORS      Apidae; Apis.
TITLE        Sawata,M., Yoshino,D., Takeuchi,H., Kamikouchi,A., Ohashi,K. and
              Kubo,T.
JOURNAL      Identification and punctate nuclear localization of a novel
              noncoding RNA, Ks-1, from the honeybee brain
PUBMED       RNA 8 (6), 772-785 (2002)
AUTHORS      Sawata,M. and Kubo,T.
JOURNAL      Direct Submission
              Submitted (05-SEP-2001) Miyuki Sawata, Department of Biological
              Sciences, Graduate School of Science, The University of Tokyo;
              Hongo 7-3-1, Bunkyo-Ku, Tokyo 113-0033, Japan
              (E-mail:m-sawata@biol.s.u-tokyo.ac.jp, Tel:81-3-5841-4448,
              Fax:81-3-5800-3553)
FEATURES     Location/Qualifiers
              1..17525
                /organism="Apis mellifera"
                /mol_type="other RNA"
                /db_xref="taxon:7460"
                /sex="female"
                /tissue_type="mushroom bodies"
                /dev_stage="adult"
                /tissue_lib="random-primed worker MB lambda-ZAP-II cDNA
                library"
                /note="consensus cDNA from contig of seven clones"
              1..17525
                /gene="Ks-1"
              1..17525
                /gene="Ks-1"
                /note="non-coding nuclear RNA"
ORIGIN
Alignment Scores:
Pred. No.:      369      Length:      17525
Score:          46.00    Matches:      8
Percent Similarity: 92.86% Conservative: 5
Best Local Similarity: 57.14% Mismatches: 1
Query Match:     52.16% Indels:      0
DB:              2      Gaps:        0

US-10-787-382-20 (1-15) x AB071139 (1-17525)

Qy      1 PheAlaValGluAsnProMetAnArgLeuValalaGluThr 14
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      7016 TTCTTATTAAAAATTGTTGATCGTTGTTGGTCACAGACT 7057

RESULT 45
AF440781      103450 bp      DNA      linear      BCT 29-AUG-2003
LOCUS        Streptomyces cinnamonensis polyether antibiotic monensin
DEFINITION   biosynthesis gene cluster, partial sequence.
ACCESSION    AF440781
VERSION      AF440781.1 GI:29122977
KEYWORDS     Streptomyces cinnamonensis
SOURCE       Streptomyces cinnamonensis
             Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
             Streptomycineae; Streptomycetaceae; Streptomyces.
REFERENCE    1 (bases 1 to 103450)
             Oliyntyk,M.
             The gene cluster for monensin biosynthesis
             Thesis (1999) University of Cambridge
             2 (bases 1 to 103450)
             Oliyntyk,M., Stark,C.B.W., Bhatt,A., Jones,M.A., Hughes-Thomas,Z.A.,
             Wilkinson,C., Oliyntyk,Z., Demydchuk,Y., Staunton,J. and
             Leadlay,P.F.
             Analysis of the biosynthetic gene cluster for the polyether
             antibiotic monensin in Streptomyces cinnamonensis and evidence for
             the role of monB and monC genes in oxidative cyclization
             Mol. Microbiol. 49 (5), 1179-1190 (2003)
             12940979

```

```

REFERENCE
AUTHORS      3 (bases 1 to 103450)
              Oliyntyk,M., Oliyntyk,Z.V. and Leadlay,P.F.
              The gene cluster for monensin biosynthesis
              Unpublished
              4 (bases 1 to 103450)
              Oliyntyk,M., Oliyntyk,Z.V. and Leadlay,P.F.
              Direct Submission
              Submitted (26-OCT-2001) Biochemistry, University of Cambridge, 80
              Tennis Court Road, Cambridge CB2 1GA, UK
FEATURES     Location/Qualifiers
              1..103450
                /organism="Streptomyces cinnamonensis"
                /mol_type="genomic DNA"
                /strain="ATCC15413"
                /db_xref="ATCC:15413"
                /db_xref="taxon:1900"
                /complement(<1..1038)
                /gene="gdhA"
                /complement(<1..1038)
                /gene="gdhA"
                /note="similar to glutamate dehydrogenase"
                /codon_start=1
                /transl_table=11
                /product="glutamate dehydrogenase"
                /protein_id="AAO65782.1"
                /db_xref="GI:29122978"
                /translation="MTTRPDTKTALSQKTSLSQTLTEIHRNPAQPEFHQAAREVLET
                LAPVIAARPEYAEAGLIERLCEPERQIVRPVMDHGRVVRNRPVFEFNSALGPYK
                GGLRPHSPVNLGVIKFLGFEQIFKNALGTGIGGKGGSDPDRGRSDAEVWRFCQSP
                MTELYRHIGHEHTDPAGDIGVGGREIGFYQYRRITNRWEAGVLTGKRNWGGSLIR
                PEATGYGNLFAAALRRERGETLEGRTAVSGSNVAIYTTQKLAALGANAVTCSDSS
                GYVDEKIGIDLLKQKVEVERAVDTVAQRGASRFVPGRRVNEVPADIALPSATQ
                NELDADATALI"
                /complement(1220..2140)
                /gene="dapA"
                /complement(1220..2140)
                /gene="dapA"
                /note="similar to dihydro-dipicolinate synthase"
                /codon_start=1
                /transl_table=11
                /product="dihydro-dipicolinate synthase"
                /protein_id="AAO65783.1"
                /db_xref="GI:29122979"
                /translation="MTLASSLEPTTPEPLNGLYVPLVTPFTDRLRLAPLARLADAE
                LSAGASGVALTGTAAATLTAEERETVIRVCSSAACRAHGAFLIVGVGNDTATATA
                LRELARGDVAALVPAPYPYRPGAGTLAHPALAEHGLPLVVDIPIYRTGQTIGA
                GRIILALRPELVVGIKHATGSDIDPTIMELLDSPFGFPAVLGGDDIVLSPLVAGAHGG
                IVASANTRTADYAEMIALWRGSAAPALGADLRLSALFTEPNPTVTKGVLHAQN
                RIPSPAVRMPLLAASADSVRRAAPLAASRK"
                2211..3152
                /note="SCN_3"
                /codon_start=1
                /transl_table=11
                /product="hypothetical DNA binding protein"
                /protein_id="AAO65784.1"
                /db_xref="GI:29122980"
                /translation="MLDVRLLHLELDRGTIAVAEALTFETASVSQQLGVLEREA
                GVPLERSGRRVLTLPAGRSVLVAHDADAVLNRLAEQVAELAGARDGGGLRIGTFPSG
                GHTIVPGLAEALSRHPALPEMVRIDSARVDSGLNAGELDALVHDIDFVPATPTT
                VDEVPILPEPMYLVTHAADTATDSGSLTAAALGPCAEPWITAKDGTGTGHAMAVRA
                CQAGFQPRIRHQVNDFTVLVAAGQAGGFVPRMAEFPAGVVLTKLPLFRRSKV
                AFRGGGAHPAIAAFVAAATTAVERMAGSRGPGAGSE"
                3264..3680
                /note="SCN_4"
                /transl_table=11
                /product="hypothetical protein"
                /protein_id="AAO65785.1"
                /db_xref="GI:29122981"
                /translation="MADDAVLELDPDRHRLGALAAVCALECTETPAVHAMIQAHEA
                SVSSEQVRLIPADAETLIPDAERLVPVLSSEALKVQECCAPQTVTDMESSELLAFRE
                TTQDQVALVHRLTAGIPQAIRALTGLDPEEIGRL"
                /complement(3684..4307)

```

```

/note="SCN_5"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AA065786.1"
/db_xref="GI:29122982"
/translation="MAVAAACAIVLPIDAVVRISAADVGVLVFFAYLLPYLAITMTVF
VSAPEQVWARREARGTFLQRYVLGTAPGGSLFIAAALVAVLWLPGLHSTTF
SALPRTLVALLVAAITCVVAVAFVTFQADNLVENERALEFPGRSPAWADYVYFAL
AAMTFGTVDVTSRDMRTVAANTVIAFVNTVAILVSLGGR"
4570..4758
/note="SCN_6"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AA065787.1"
/db_xref="GI:29122983"
/translation="MTWMDKLKQMLKHEDKAGQIDIKAGDFVDGKTQYSGQVDTA
QDKLRDQFGSDQEPQPR"
5058..5612
/note="SCN_7"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AA065788.1"
/db_xref="GI:29122984"
/translation="MGTAQSORQAAAPGCAAFVRFVLCGGVGLASSFAVVVALASVW
PWALANLVAVTVVATTELHARFTFGAGGRATWFOHAQSAGSAAAVAVTCVAMFVL
QQLVAPGAVLEQVYLSASALAGVAFVVLVLFVARNRSLPAAAVARTAPVRVP
APVATVAHAASRPAGPAALCPAA"
complement(5693..6010)
/note="SCN_8"
/codon_start=1
/transl_table=11
/product="acyl carrier protein"
/protein_id="AA065789.1"
/db_xref="GI:29122985"
/translation="MTSDHTSGQDATELEKQLAATPEEREKLLTDIRTOAGTLLN
TTLSDSNFLENGLSLTALELTKLMTLTGMEIAWVAIVENPTPAQLAHLGQELAH
TTA"
complement(6045..8531)
/note="SCN_9"
/codon_start=1
/transl_table=11
/product="ketosynthase-like protein"
/protein_id="AA065790.1"
/db_xref="GI:29122986"
/translation="WANEKLVYLLKWTAEHLHQOQLRELKAAQHBPAAVSVMACR
LFGKTRTPDDLWLVSEGRDAVTGFPDRAWELPEERPYAELGGFLDDAGDGFDD
ISDTAVATEPLQRLMLHAWETVERGHIAPIHLASTLTGTVVYGATGHDYARTETAP
DELLPYLGGTSGSLVSGRIAVALGLEGPAISVDTACSSSLVALHLAQQALRRGECGL
ALAGGTVNSTHTTHAFHQSLAQDGRCKPFAAADGMGLGEGVGLVLLERLGDAR
KNGHPLAVIRGAVNQDAGYGLAPNGPSQOHVIRAAADAGLTPQDIDVAEAGHT
LTPGIDAEVALLATYGADSPDRPLMLGVSQKNTGHTQGAAGAAALIKMVQAFRHG
GTPIIDAEVSTVPLAAKKGAVLLTEAVDWPREEPRRVGISAFATSGTNWHLILEE
PVDVAPDPAARDOTSPVAPELPVAMSLSARTPEALRAQAKALVTHLAATDPAFSPA
EYVLSIATRSPLERAVLTGTDHTELLAAAPALAGSDHDPDLVSTPCAGPKLTAWH
FDRPADGVTTGAAGFAGCATGATFGAAGFAGFASAPLFPASAFPEARALLTHL
PTPLTPHSELARFAVHTALRLLELTGVRPHLTGDDGVGHIAAAYAGILFLDDACR
LAAHAAQAAGAEQGPAPPDAYEPVLTQLTFRATLTLTSTAPADTPIASADYWHH
LTSAPATPTPTHTLHLGALSPEGTQTSFVSAVLLTALARLHTTGGTVDWTPLVRR
TPHRTIDLPVTSFOATRYMLHDHTAAAV"
complement(8643..9542)
/note="monCII"
complement(8643..9542)
/note="monCII"
/note="monensin epoxide hydrolase; MonCII"
/codon_start=1
/transl_table=11
/product="monensin cyclase"
/protein_id="AA065791.1"
/db_xref="GI:29122987"
/translation="MKNLRIPVSVTVSLNVRYPADPGAPGRPFLLHLHGLSNARMW
```

gene

CDS

gene

CDS

CDS

CDS

gene

CDS

```

DEVARLAAAGHPAYAVDHRHGSESTPPPGYDNATVVTDLVAAVTALDLSGALVAGH
SWGHALRLAAEHPDLVAGLALIDGGWYEFDFGVKRAFWEKRTADVVRRAQQGTTSA
DMRAYLRATHPDWSPTIEARLADYRVGPDGLIPLRTSTQVMSIVAGLQREAPADWY
PKVTVPVRLPLPAIPQLSQVRAWVAEEAALEQVSVRWYPGSDHDLHAGAPDEIA
ADLILLARSCEAMEPGKAGVRPA"
complement(9596..10426)
/note="monE"
/note="monE"
/note="SAM dependent; MonE"
/codon_start=1
/transl_table=11
/product="monensin 3-O-methyl transferase"
/protein_id="AA065792.1"
/db_xref="GI:29122988"
/translation="MNKTVAPESDIGHYDGHKVFDMLTQLDGNLHYGYWFDGGEQQ
ATPDEAMVQMTDEMIRLDDPAGRVLDIGCGNGTAPMQLARARDEVEVGI SVSARQV
ERGNRAREAGLADRVRFEOVDAMNLPDFDGSFDHCWALESMHMPKQOVLTSARHV
VKGARPMIADMYVLNPDFRPRPTVSDTTIYAALTDIGDYDPLIFRAGWTVLELTD
ITRETAKTYDGYVEWIRAHRDVEYVDIIGVEGYELFLHNQAALGKMPGLGYIFATAQRP
"
10656..12191
/note="monT"
/note="monT"
10656..12191
/note="monI"
/note="membrane transporter; putative monensin
transporter; MonT"
/codon_start=1
/transl_table=11
/product="putative monensin resistance protein"
/protein_id="AA065793.1"
```

Alignment Scores:

Pred. No.:	2.9e+03	Length:	103450
Score:	46.00	Matches:	9
Percent Similarity:	84.62%	Conservative:	2
Best Local Similarity:	69.23%	Mismatches:	2
Query Match:	62.16%	Indels:	0
DB:	1	Gaps:	0

US-10-787-382-20 (1-15) x AF440781 (1-103450)

Qy 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGlu 13

Db 98150 TTCTCTGTGCACGACCGCATGTCAGTCTGCTGCCGAA 98188

RESULT 46

CP000088_30/c

WPCOMMENT

Sequence split into 37 fragments LOCUS CP000088 Accession CP000088

Fragment Name Begin End

CP000088_00	1	110000
CP000088_01	100001	210000
CP000088_02	200001	310000
CP000088_03	300001	410000
CP000088_04	400001	510000
CP000088_05	500001	610000
CP000088_06	600001	710000
CP000088_07	700001	810000
CP000088_08	800001	910000
CP000088_09	900001	1010000
CP000088_10	1000001	1110000
CP000088_11	1100001	1210000
CP000088_12	1200001	1310000
CP000088_13	1300001	1410000
CP000088_14	1400001	1510000
CP000088_15	1500001	1610000
CP000088_16	1600001	1710000
CP000088_17	1700001	1810000
CP000088_18	1800001	1910000
CP000088_19	1900001	2010000
CP000088_20	2000001	2110000
CP000088_21	2100001	2210000

```
CP000088_22 2200001 2310000
CP000088_23 2300001 2410000
CP000088_24 2400001 2510000
CP000088_25 2500001 2610000
CP000088_26 2600001 2710000
CP000088_27 2700001 2810000
CP000088_28 2800001 2910000
CP000088_29 2900001 3010000
CP000088_30 3000001 3110000
CP000088_31 3100001 3210000
CP000088_32 3200001 3310000
CP000088_33 3300001 3410000
CP000088_34 3400001 3510000
CP000088_35 3500001 3610000
CP000088_36 3600001 3642249
Continuation (31 of 37) of CP000088 from base 3000001 (CP000088 Thermobifida fusca YX, C

Alignment Scores:
Pred. No.: 3,11e+03 Length: 110000
Score: 46.00 Matches: 9
Percent Similarity: 85.71% Conservative: 3
Best Local Similarity: 64.29% Mismatches: 2
Query Match: 62.16% Indels: 0
DB: 1 Gaps: 0

US-10-787-382-20 (1-15) x CP000088_30 (1-110000)

QY 2 AlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
||||| :|||:|||||:|||||:|||||
Db 18616 GCGTCCACCTCCCGTGAACAGGGTGTGCGCGACATTG 18575

RESULT 47
AC007086/c
LOCUS AC007086 186241 bp DNA linear INV 23-MAR-2001
DEFINITION Drosophila melanogaster, chromosome 2R, region 45A-46A, BAC clone
ACCESSION AC007086
VERSION AC007086.10 GI:13435225
KEYWORDS HTG.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
REFERENCE 1 (bases 1 to 186241)
AUTHORS Celniker,S.E., Adams,M.D., Kronmiller,B., Tyler,D., Wan,K.H.,
Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Brandon,R.C.,
Rogers,Y., An,H., Baldwin,D., Banzon,J., Beeson,K.Y., Busan,D.A.,
Carlson,J.W., Center,A., Champe,M., Davenport,L.B., Dietz,S.M.,
Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dreanek,D., Farfan,D.,
Ferrier,S., Frise,E., Galle,R.F., Garg,N.S., George,R.A.,
Gonzalez,M., Houck,J., Hoskins,R.A., Hostin,D., Howland,T.J.,
Ibegwam,C., Jalali,M., Kruse,D., Li,P., Mattel,B., Moshrefi,A.,
McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J.,
Pacלב,J., Paragas,V., Park,S., Patel,S., Pfeiffer,B.,
Phouanavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F.,
Shaplenavong,S., Strong,R., Svirskas,R., Tector,C., Williams,S.M.,
Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,J.C.
Sequencing of Drosophila chromosome 2R, region 45A-46A
Unpublished
2 (bases 1 to 186241)
AUTHORS Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazek,R.G.,
Buenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
Doyle,C.M., Farfan,D.B., Galle,R., George,R.A., Harris,N.L.,
Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P.,
Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacלב,J.M., Park,S.,
Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Snir,E.,
Svirskas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and
Rubin,G.M.
Direct Submission
TITLE Drosophila melanogaster, chromosome 2R, region 45A-46A, BAC clone
JOURNAL Submitted (16-MAR-1999) Drosophila Genome Center, Lawrence Berkeley
Laboratory, MS 64-121, Berkeley, CA 94720, USA
```

```
COMMENT On Mar 23, 2001 this sequence version replaced gi:6957981.
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory, MS 64-121
Berkeley, CA 94720
This sequence was assembled using end sequences from a whole genome
shotgun and from subclones of this BAC and its neighboring clones.
For further information about this sequence, including its location
and relationship to other sequences, please visit our sequence
archive Web site (http://www.fruitfly.org/sequence/) or send email
to bdgpf@fruitfly.berkeley.edu.
FEATURES
Location/Qualifiers
1..186241
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/strain="Y; cn bw sp"
/db_xref="taxon:7227"
/chromosome="2R"
/map="45A-46A"
/clone="BACR14J24 (D585)"
/clone_lib="RPCI-98 (Roswell Park Cancer Institute
Drosophila melanogaster BAC library, partial EcoRI in
pBACE3.6)"
ORIGIN
Alignment Scores:
Pred. No.: 5,73e+03 Length: 186241
Score: 46.00 Matches: 8
Percent Similarity: 83.33% Conservative: 2
Best Local Similarity: 66.67% Mismatches: 2
Query Match: 62.16% Indels: 0
DB: 2 Gaps: 0

US-10-787-382-20 (1-15) x AC007086 (1-186241)

QY 1 PheAlaValGluAsnProMetAsnArgLeuValAla 12
||||| :|||:|||||:|||||:|||||
Db 89251 TTGTGCTGCTGCAACCCATTAAACGATGTCGCG 89216

RESULT 48
AC022809/c
LOCUS AC022809 201020 bp DNA linear PRI 06-MAY-2002
DEFINITION Homo sapiens chromosome 18, clone RP11-699A5, complete sequence.
ACCESSION AC022809
VERSION AC022809.6 GI:20429562
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 201020)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 18, clone RP11-699A5
JOURNAL Unpublished
2 (bases 1 to 201020)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Becker,L., Beda,F.,
Boguslavsky,L., Bouckgaiter,B., Brown,A., Burkett,G., Castle,A.,
Chospel,Y., Collangelo,M., Collins,S., Collymore,A., Cooke,P.,
DeArelano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
Ferreira,P., Fitzhugh,W., Forrest,C., Gage,D., Galagan,J.,
Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lander,B., Lechoczky,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., Olivari,T.M., Peterson,K.,
Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
Roy,A., Santos,R., Severi,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
Zimmer,A. and Zody,M.
```



```

/rpt family="AT_rich"
complement(21659..21739)
/rpt family="U6"
21783..22084
/rpt family="AluX"
22189..22189
/rpt family="AT_rich"
22563..22605
/rpt family="AT_rich"
complement(22650..22941)
/rpt family="AluJo"
23615..23837
/rpt family="MER104"
23838..24137
/rpt family="Aluub"
24138..24334
/rpt family="MER113"
24459..24760
/rpt family="AluX"
complement(25531..25777)

Alignment Scores:
Pred. No.: 6.26e+03 Length: 201020
Score: 46.00 Matches: 7
Percent Similarity: 83.33% Conservative: 3
Best Local Similarity: 58.33% Mismatches: 2
Query Match: 62.16% Indels: 0
DB: 8 Gaps: 0

US-10-787-382-20 (1-15) x AC022809 (1-201020)

Qy 3 ValGluAenProMetAsnArgLeuValAlaGluThr 14
Db 56734 ATAGAAACCCATTTATGAAATAATAGCTGAACACT 56699

RESULT 49
AC164180 linear HTG 01-JUL-2005
LOCUS Bos taurus clone CH240-154E1, *** SEQUENCING IN PROGRESS ***, 27
DEFINITION unordered pieces.
AC164180
AC164180.3 GI:68300674
HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
KEYWORDS Bos taurus (cow)
SOURCE
ORGANISM

REFERENCE
AUTHORS Muzny, D.M., Metaker, M., Lee, A., Adams, C., Alder, J.,
Allen, C., Allen, H., Albrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escoto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebregorgis, B., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, W.,
Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S.B., Hodgson, A., Hogues, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,

```


Lorensuhsuwa, L., Loulsegged, H., Lozado, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mamoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwokedemen, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puaso, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodley, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajs, D., Sneed, A., Shvartsbeyn, A., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.
 Direct Submission
 Unpublished
 2 (bases 1 to 249287)
 Worley, K. C.
 Direct Submission
 Submitted (21-MAY-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 249287)
 Cow Genome Sequencing Consortium.
 Direct Submission
 Submitted (01-JUL-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: FGOM
 Center clone name: CH240-105K16
 ----- Summary Statistics
 Assembly program: Atlas 3.0;
 Consensus quality: 236931 bases at least Q40
 Consensus quality: 239534 bases at least Q30
 Consensus quality: 241569 bases at least Q20
 Estimated insert size: 244133; sum-of-contigs estimation
 Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 32 contigs. The true order of the pieces
 * is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence.
 * as soon as it is available and the accession number will
 * be preserved.
 *
 * 1
 * 2228: contig of 2228 bp in length
 * 2278: gap of 50 bp
 * 2279: contig of 4389 bp in length
 * 6667: gap of unknown length
 * 6767: gap of 1768 bp in length
 * 8535: contig of 398 bp
 * 8934: contig of 3891 bp in length
 * 12824: contig of 727 bp
 * 13551: gap of 2189 bp in length
 * 15740: contig of 501 bp
 * 15741: gap of 501 bp
 * 16241: gap of 501 bp
 * 16242: contig of 5024 bp in length
 * 21266: contig of 450 bp
 * 21715: gap of 450 bp
 * 21716: contig of 9593 bp in length
 * 31308: contig of 50 bp
 * 31309: gap of 50 bp
 * 3534: contig of 4576 bp in length
 * 35935: gap of 50 bp
 * 35936: gap of 50 bp
 * 35985: contig of 7433 bp in length
 * 43417: gap of 50 bp
 * 43418: contig of 3346 bp in length
 * 46813: gap of 50 bp
 * 46814: gap of 50 bp
 * 56291: contig of 9428 bp in length
 * 56341: gap of 50 bp
 * 56292: gap of 50 bp
 * 56342: contig of 5698 bp in length
 * 62089: gap of 50 bp
 * 62090: contig of 2219 bp in length
 * 64309: gap of 684 bp
 * 64309: contig of 32975 bp in length
 * 97967: gap of 50 bp
 * 97968: gap of 50 bp
 * 98017: gap of 50 bp
 * 100110: contig of 2093 bp in length
 * 100111: gap of 50 bp
 * 100160: gap of 50 bp
 * 109515: contig of 9355 bp in length
 * 109516: gap of 50 bp
 * 133886: contig of 24321 bp in length
 * 133887: gap of 180 bp
 * 134067: contig of 5605 bp in length
 * 139671: gap of 50 bp
 * 139722: contig of 1255 bp in length
 * 140977: gap of 50 bp
 * 141026: gap of 50 bp
 * 143027: contig of 2582 bp in length
 * 143609: gap of 50 bp
 * 147495: contig of 3837 bp in length
 * 147496: gap of 50 bp
 * 147546: contig of 8068 bp in length
 * 155613: gap of 50 bp
 * 155614: gap of 50 bp
 * 158869: contig of 3206 bp in length
 * 158870: gap of 50 bp
 * 158920: contig of 4199 bp in length
 * 163119: gap of 50 bp
 * 163169: contig of 47028 bp in length
 * 210197: gap of 50 bp
 * 210247: gap of 50 bp
 * 225457: contig of 15210 bp in length
 * 225814: gap of 357 bp
 * 238084: contig of 12270 bp in length
 * 238183: gap of unknown length
 * 239184: contig of 1000 bp in length
 * 239284: gap of unknown length
 * 241097: contig of 1813 bp in length
 * 241197: gap of unknown length
 * 242602: contig of 1405 bp in length
 * 242702: gap of unknown length
 * 245413: contig of 2711 bp in length
 * 245512: gap of unknown length
 * 249287: contig of 3775 bp in length.
 *
 * Location/Qualifiers
 1. 249287
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /db_xref="taxon:9913"

FEATURES
source

```

gap
/clone="CH240-105K16"
2229. .2278
/estimated_length=50
6668. .6767
/estimated_length=unknown
8536. .8933
/estimated_length=398
12825. .13551
/estimated_length=727
15741. .16241
/estimated_length=501
21266. .21715
/estimated_length=450
31309. .31358
/estimated_length=50
35935. .35984
/estimated_length=50
43418. .43467

```

Alignment Scores:

Pred. No.:	8.046+03	Length:	249287
Score:	46.00	Matches:	8
Percent Similarity:	92.31%	Conservative:	4
Best Local Similarity:	61.54%	Mismatches:	1
Query Match:	62.16%	Indels:	0
DB:	14	Gaps:	0

US-10-787-382-20 (1-15) x AC161834 (1-249287)

QY 3 ValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
 DB 233045 CTGGAAATCCCATGGCAGAGTGGTGGCTCAGACGGTA 233007

Search completed: December 21, 2005, 18:08:13
 Job time : 3135 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 21, 2005, 14:31:31 ; Search time 401 Seconds

(without alignments)
249.303 Million cell updates/sec

Title: US-10-787-382-20

Perfect score: 74

Sequence: 1 PAVENPMRLVAETL 15

Scoring table:

BGAPop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2.1/USPTO.spool/US10787382/runat.16122005.132436.8002/app.query.fasta.1.199
-DB=N_Geneseq -QFWT=fastap -SUPFIX=p2n.rng -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=DIOSUM62 -TRANS=human40.cdi -LIST=500
-DOCALIGN=200 -THR SCORE=ptc -THR MAX=100 -THR MIN=0 -ALIGN=50 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10787382 @CNC 1.1.727 @runat.16122005.132436.8002 -NCPU=6 -ICPU=3
-NO.MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: Geneseq 21.*
1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001s.*
5: Geneseq2001bs.*
6: Geneseq2002as.*
7: Geneseq2002bs.*
8: Geneseq2003as.*
9: Geneseq2003bs.*
10: Geneseq2003cs.*
11: Geneseq2003ds.*
12: Geneseq2004as.*
13: Geneseq2004bs.*
14: Geneseq2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	74	100.0	252	4	Aaf74305 Canine in
2	74	100.0	345	3	Aaz55550 Canine ma
3	74	100.0	345	3	Aaz55551 Canine ma
4	74	100.0	402	3	Aaz55548 Canine in

c	5	74	100.0	402	3	AAZ55549	Canine in
	6	74	100.0	405	4	AAF74300	Canine in
	7	74	100.0	610	3	AAZ55546	Canine in
c	8	74	100.0	610	3	AAZ55547	Canine in
	9	60	81.1	838	3	AAZ44265	Porcine I
	10	55	74.3	399	2	AAT50756	Ovine II-
	11	55	74.3	520	2	AAT50755	Ovine II-
	12	46	62.2	103599	4	ABX04971	S. cinna
c	13	46	62.2	312477	12	ADP69744	Pepper m
	14	43	58.1	673	3	ABA97721	Human ROC
	15	43	58.1	2256	10	ADE54216	Human gen
c	16	42	56.8	363	5	RAF67150	Novel hum
c	17	42	56.8	507	6	ABQ91289	M. capsul
c	18	42	56.8	14707	6	AAZ53529	Genomic D
	19	42	56.8	73771	11	ACN44938	Human gen
c	20	41	55.4	238	10	ADE06666	Hepatit
	21	41	55.4	250	10	ADE06664	Hepatit
c	22	41	55.4	381	10	ACF68089	Photorhab
	23	41	55.4	924	12	ADO07799	Fly polyn
	24	41	55.4	941	4	ABL30155	Drosophil
	25	41	55.4	1545	13	ADT42017	Bacterial
c	26	41	55.4	1798	6	ABZ16494	Arabidops
	27	41	55.4	2000	10	ACC61263	Gene sequ
c	28	41	55.4	2000	10	ADK63721	Disease t
	29	41	55.4	2442	10	ADE06713	ORF1 gene
	30	41	55.4	2941	4	ABL30154	Drosophil
c	31	41	55.4	3004	10	ADD93578	Arabidops
	32	41	55.4	5138	10	ADE06707	ORF1 gene
c	33	41	55.4	6496	4	ABL30196	Drosophil
	34	41	55.4	7230	10	ADE06697	Hepatit
c	35	41	55.4	7233	10	ADE06694	Hepatit
c	36	41	55.4	10708	13	ADV41357	Rat cardi
	37	41	55.4	110000	10	ACF67367	Continuati
	38	41	55.4	110000	10	ACF67367	Continuati
	39	41	55.4	110000	10	ACF65384	Continuati
	40	41	55.4	110000	10	ACF65384	Continuati
c	41	40	54.1	459	4	ABA26715	Probe #51
	42	40	54.1	480	8	ACA43050	Prokaryot
c	43	40	54.1	567	4	ABA31096	Probe #95
	44	40	54.1	569	12	ACH72974	Human gen
c	45	40	54.1	913	5	AAZ73628	DNA encod
	46	40	54.1	2310	12	ADN12289	Nematode
c	47	40	54.1	2518	5	AAZ73627	DNA encod
	48	40	54.1	2520	5	AAZ73629	DNA encod
c	49	40	54.1	2591	11	ADM01809	Human CDN
	50	40	54.1	6692	4	ABL14942	Drosophil
c	51	40	54.1	6692	4	ABL20236	Drosophil
	52	40	54.1	103471	12	ADQ97668	Mouse can
	53	40	54.1	110000	6	ABA03041	Continuati
	54	40	54.1	110000	13	ABD32791	Continuati
c	55	40	54.1	170834	10	AAD62833	Human BAC
	56	39	52.7	71	2	AAT71466	Glioblast
c	57	39	52.7	121	11	ADZ42577	Human gen
	58	39	52.7	563	9	ACH39654	Human toe
	59	39	52.7	673	12	ADQ17913	Human sof
	60	39	52.7	806	4	AAF17601	Human bre
	61	39	52.7	806	4	AAF17587	Human bre
	62	39	52.7	806	4	AAZ47017	Human bre
	63	39	52.7	806	4	AAZ47031	Human bre
	64	39	52.7	806	6	ABK95066	Human bre
	65	39	52.7	806	6	ABT08672	Human bre
	66	39	52.7	806	6	ABT08686	Human bre
	67	39	52.7	806	6	ABS63618	Human bre
	68	39	52.7	806	6	ABS63632	Human bre
	69	39	52.7	806	10	ABT32830	Human tum
	70	39	52.7	806	10	ABT32844	Human tum
	71	39	52.7	806	11	ADL92734	Human bre
	72	39	52.7	806	11	ADL92748	Human bre
	73	39	52.7	806	12	ADE44024	Human CDN
	74	39	52.7	806	12	ADE44038	Human CDN
	75	39	52.7	934	12	ADM16745	PERS-P1-H
c	76	39	52.7	1171	4	AAI71121	Escherich
	77	39	52.7	1171	4	AAI71102	Shigella

224	37	50.0	531	13	ACN62188	Acn62188 Cotton gy	c 297	37	50.0	1906	4	ABA89379	Abas89379 Escherich
225	37	50.0	543	8	ABZ54155	ABZ54155 Aspergill	298	37	50.0	1945	10	ABD53890	ABD53890 Primary r
226	37	50.0	555	4	AKX92641	AKX92641 Human cDN	299	37	50.0	1976	5	AAF93822	AAF93822 Human cDN
227	37	50.0	555	12	ADL29068	ADL29068 3' end of	300	37	50.0	1976	14	ADY63154	ADY63154 Human clo
228	37	50.0	576	6	ABZ14403	ABZ14403 Arabidops	301	37	50.0	1989	13	ADS64016	ADS64016 Bacterial
229	37	50.0	576	8	ADA68274	ADA68274 Arabidops	302	37	50.0	1989	13	ADS63641	ADS63641 Bacterial
230	37	50.0	601	13	ACNS2466	ACNS2466 Cotton an	c 303	37	50.0	2000	8	ADA72109	ADA72109 Rice gene
231	37	50.0	603	5	AAF94164	AAF94164 Primer sp	304	37	50.0	2000	8	ADA72674	ADA72674 Rice gene
232	37	50.0	611	14	ADY63591	ADY63591 Human clo	305	37	50.0	2000	8	ADA71656	ADA71656 Rice gene
233	37	50.0	611	8	ABZ54724	ABZ54724 Aspergill	306	37	50.0	2000	11	ACL38104	ACL38104 Rice stre
234	37	50.0	615	13	ADS47911	ADS47911 Bacterial	307	37	50.0	2000	11	ACL37113	ACL37113 Rice stre
235	37	50.0	634	4	ABA89381	ABA89381 Escherich	308	37	50.0	2000	12	ADJ40987	ADJ40987 Plant cDN
236	37	50.0	693	8	ACA24591	ACA24591 Prokaryot	c 309	37	50.0	2024	6	ABQ76378	ABQ76378 S. cerevi
237	37	50.0	831	14	ACL72098	ACL72098 M. xanthu	c 310	37	50.0	2024	12	ADC24593	ADC24593 Human sof
238	37	50.0	833	2	AAQ25034	AAQ25034 Bovine RS	311	37	50.0	2199	10	ADC08389	ADC08389 Rice DNA
239	37	50.0	835	3	AAF09241	AAF09241 Fusarium	312	37	50.0	2206	14	ADV97739	ADV97739 cDNA sequ
240	37	50.0	835	13	ADU53282	ADU53282 Fusarium	313	37	50.0	2231	4	AAH99815	AAH99815 Human pro
241	37	50.0	835	14	ADZ91285	ADZ91285 Fusarium	314	37	50.0	2245	4	AAK82349	AAK82349 Human imm
242	37	50.0	936	6	ABK73208	ABK73208 Bacillus	315	37	50.0	2283	4	ABL20647	ABL20647 Drosophil
243	37	50.0	936	8	ACA32419	ACA32419 Prokaryot	c 316	37	50.0	2328	10	ADG45274	ADG45274 Mutant De
244	37	50.0	941	8	ACA49138	ACA49138 Prokaryot	c 317	37	50.0	2328	10	ADG45273	ADG45273 Mutant De
245	37	50.0	963	5	AAH67228	AAH67228 C. glutami	c 318	37	50.0	2328	10	ADG45275	ADG45275 Mutant De
246	37	50.0	1000	3	AAK57826	AAK57826 Arachidon	c 319	37	50.0	2328	10	ADG45276	ADG45276 Mutant De
247	37	50.0	1032	2	AAK94209	AAK94209 Bacillus	c 320	37	50.0	2328	10	ADG45277	ADG45277 Mutant De
248	37	50.0	1041	3	AAZ51712	AAZ51712 Burkholde	c 321	37	50.0	2328	10	ADG45278	ADG45278 Mutant De
249	37	50.0	1041	3	AAZ51713	AAZ51713 Burkholde	c 322	37	50.0	2328	10	ADG45272	ADG45272 Mutant De
250	37	50.0	1041	3	AAZ51714	AAZ51714 Burkholde	c 323	37	50.0	2328	12	ADG45271	ADG45271 DNA encod
251	37	50.0	1053	3	AAZ51714	AAZ51714 Burkholde	c 324	37	50.0	2328	12	ADN75087	ADN75087 Wild-type
252	37	50.0	1068	8	ACA23794	ACA23794 Prokaryot	c 325	37	50.0	2328	12	ADN59949	ADN59949 Pyrococcu
253	37	50.0	1068	8	ACA23794	ACA23794 Prokaryot	c 326	37	50.0	2328	12	ADQ26670	ADQ26670 Mutant De
254	37	50.0	1134	5	AAK589404	AAK589404 DNA encod	c 327	37	50.0	2328	12	ADQ26669	ADQ26669 Mutant De
255	37	50.0	1149	3	AAZ53806	AAZ53806 Neisseria	c 328	37	50.0	2328	13	ADQ7146	ADQ7146 Pfu DNA p
256	37	50.0	1149	3	AAZ53806	AAZ53806 Neisseria	c 329	37	50.0	2328	14	ADY80957	ADY80957 Mutant De
257	37	50.0	1149	13	ADU52221	ADU52221 Fusarium	c 330	37	50.0	2331	2	AAQ50112	AAQ50112 Gene enco
258	37	50.0	1149	13	ADU52221	ADU52221 Fusarium	c 331	37	50.0	2331	2	AAQ50112	AAQ50112 Gene enco
259	37	50.0	1225	14	ADZ90224	ADZ90224 Fusarium	c 332	37	50.0	2333	4	ABL27378	ABL27378 Drosophil
260	37	50.0	1273	14	ADW16954	ADW16954 Pinus rad	c 333	37	50.0	2493	10	ABZ39948	ABZ39948 N. gonorr
261	37	50.0	1275	14	AEK67684	AEK67684 Rice geno	c 334	37	50.0	2517	13	ADT07246	ADT07246 Pfu DNA p
262	37	50.0	1329	13	ADT46354	ADT46354 Bacterial	c 335	37	50.0	2517	13	ADT07245	ADT07245 Pfu DNA p
263	37	50.0	1433	2	AAQ65397	AAQ65397 Elm C10:O	c 336	37	50.0	2526	13	ADT07230	ADT07230 Pfu DNA p
264	37	50.0	1433	2	AAQ92306	AAQ92306 Elm class	c 337	37	50.0	2526	13	ADT07229	ADT07229 Pfu DNA p
265	37	50.0	1433	2	AAK36108	AAK36108 Elm acyl-	c 338	37	50.0	2532	13	ADT07186	ADT07186 Pfu DNA p
266	37	50.0	1433	2	AAV15230	AAV15230 C10:0-ACP	c 339	37	50.0	2535	13	ADR25649	ADR25649 Breast ca
267	37	50.0	1484	3	AAK47540	AAK47540 Arabidops	c 340	37	50.0	2535	13	ADT07187	ADT07187 Pfu DNA p
268	37	50.0	1486	3	AAK40456	AAK40456 Arabidops	c 341	37	50.0	2564	5	AAK89190	AAK89190 DNA encod
269	37	50.0	1521	11	ADQ02990	ADQ02990 Pseudomon	c 342	37	50.0	2583	10	ADB62856	ADB62856 Human cDN
270	37	50.0	1524	9	AAK59829	AAK59829 Yeast cys	c 343	37	50.0	2655	4	ABL13062	ABL13062 Drosophil
271	37	50.0	1524	10	ACC61802	ACC61802 Gene sequ	c 344	37	50.0	2724	5	ABA15116	ABA15116 Human per
272	37	50.0	1524	10	ADK62743	ADK62743 Disease t	c 345	37	50.0	2864	4	AAK94275	AAK94275 Human ful
273	37	50.0	1531	12	ADO62068	ADO62068 Transcrip	c 346	37	50.0	2864	12	ADL30875	ADL30875 Full leng
274	37	50.0	1533	1	AAK82431	AAK82431 B cell di	c 347	37	50.0	2955	5	AAH65842	AAH65842 C. glutami
275	37	50.0	1534	2	AAK88013	AAK88013 Murine in	c 348	37	50.0	2955	8	ACA00203	ACA00203 C. glutam
276	37	50.0	1596	4	AAH34605	AAH34605 Human col	c 349	37	50.0	3062	3	AAZ87684	AAZ87684 Human G p
277	37	50.0	1602	14	ABE28967	ABE28967 Pinus rad	c 350	37	50.0	3075	13	ADT07208	ADT07208 Pfu DNA p
278	37	50.0	1620	4	AAK71697	AAK71697 Human imm	c 351	37	50.0	3075	13	ADT07209	ADT07209 Pfu DNA p
279	37	50.0	1620	8	ACA19515	ACA19515 Pseudomon	c 352	37	50.0	3106	2	AAV70135	AAV70135 Pyruvate
280	37	50.0	1623	2	AAK14925	AAK14925 T cell re	c 353	37	50.0	3106	4	AAK723625	AAK723625 TufD codi
281	37	50.0	1623	2	AAV64062	AAV64062 Plasmid p	c 354	37	50.0	3316	10	ADI02625	ADI02625 Human cDN
282	37	50.0	1656	14	AEA10860	AEA10860 Magnaport	c 355	37	50.0	3360	2	AAK07285	AAK07285 Human len
283	37	50.0	1656	14	AEA12337	AEA12337 DNA encod	c 356	37	50.0	3370	3	AAA16630	AAA16630 Human sec
284	37	50.0	1656	14	AEA14792	AEA14792 Sub-famil	c 357	37	50.0	3377	12	ADQ18296	ADQ18296 Human sof
285	37	50.0	1674	11	ABD03041	ABD03041 Pseudomon	c 358	37	50.0	3377	13	ADP24259	ADP24259 PRO polyp
286	37	50.0	1682	4	AAK71694	AAK71694 Human imm	c 359	37	50.0	3389	10	ADP79882	ADP79882 Human put
287	37	50.0	1682	4	AAK71694	AAK71694 Human imm	c 360	37	50.0	3389	11	ADP65829	ADP65829 Human end
288	37	50.0	1682	4	AAK71694	AAK71694 Human imm	c 361	37	50.0	3389	11	ADP65751	ADP65751 Human SM-
289	37	50.0	1761	3	AAK71697	AAK71697 Human imm	c 362	37	50.0	3420	2	AAQ43965	AAQ43965 Pyrococcu
290	37	50.0	1761	3	AAK71697	AAK71697 Human imm	c 363	37	50.0	3420	2	AAQ43518	AAQ43518 Pyrococcu
291	37	50.0	1812	5	ABD02924	ABD02924 Pseudomon	c 364	37	50.0	3420	2	AAK17981	AAK17981 Pyrococcu
292	37	50.0	1812	5	AAK593842	AAK593842 DNA encod	c 365	37	50.0	3465	5	AAH67229	AAH67229 C. glutami
293	37	50.0	1821	4	ABL24609	ABL24609 Drosophil	c 366	37	50.0	3465	14	ABE13128	ABE13128 C. glutam
294	37	50.0	1851	13	ADT41607	ADT41607 Bacterial	c 367	37	50.0	3508	5	ADL63625	ADL63625 Human ova
295	37	50.0	1861	6	ABK75009	ABK75009 Bacillus	c 368	37	50.0	3575	12	ADQ87654	ADQ87654 Human tum
296	37	50.0	1887	8	ACA00696	ACA00696 C. glutam	c 369	37	50.0	3575	13	ADQ84415	ADQ84415 Human tum
	37	50.0	1892	4	ABL20645	ABL20645 Drosophil		37	50.0	3595	10	ADJ87396	ADJ87396 DNA repli

370	37	50.0	3659	12	ADQ22831	Human sof
371	37	50.0	4024	11	ACN8855	Breast ca
372	37	50.0	4056	4	ABL09821	Drosophila
373	37	50.0	4133	13	ACN42874	Human dia
374	37	50.0	4158	10	ADG77045	Human nuc
375	37	50.0	4173	13	ADG92941	Mitogen-a
376	37	50.0	4190	4	ABL24608	Drosophila
377	37	50.0	4278	8	ACN33304	Prokaryot
378	37	50.0	4320	10	ADH84843	Enterococ
379	37	50.0	4437	6	AA316840	Rat CIRL
380	37	50.0	4464	12	ADM97581	Human cal
381	37	50.0	4497	10	ADC26274	Human NOV
382	37	50.0	4571	13	ADS10070	Human the
383	37	50.0	4675	8	AA049443	Human kin
384	37	50.0	4687	12	ADQ24075	Human sof
385	37	50.0	4707	2	AA770813	Deep Vent
386	37	50.0	4707	2	AAV68142	DNA polym
387	37	50.0	4707	13	ADS11410	Human the
388	37	50.0	4780	11	ACN44533	Mouse mRN
389	37	50.0	4828	2	AAV70138	Pyruvate
390	37	50.0	4928	4	AA723627	TutD/8 ge
391	37	50.0	4983	4	ABL03779	Drosophila
392	37	50.0	5008	8	ABX63351	Human cdn
393	37	50.0	5362	4	ABL04953	Drosophila
394	37	50.0	5439	6	ABL33402	Human imm
395	37	50.0	5602	10	ABD58940	Toxicity-
396	37	50.0	5602	10	ABD53676	Primary r
397	37	50.0	5656	4	ABL10071	Drosophila
398	37	50.0	5758	4	ABL23284	Drosophila
399	37	50.0	5999	4	ABL13590	Drosophila
400	37	50.0	6279	8	ABT18891	Aspergill
401	37	50.0	6279	8	ABT20711	Aspergill
402	37	50.0	6330	8	ABT18297	Aspergill
403	37	50.0	6330	8	ABT20113	Aspergill
404	37	50.0	6727	2	AA788014	Murine IL
405	37	50.0	6727	3	AA73648	Murine IL
406	37	50.0	6727	8	ABX04302	Mouse Int
407	37	50.0	6727	12	ADR11979	Murine in
408	37	50.0	7223	4	ABL03778	Drosophila
409	37	50.0	8330	8	ABT19517	Aspergill
410	37	50.0	8330	8	ABT17703	Aspergill
411	37	50.0	8614	4	AAK74343	Human imm
412	37	50.0	8966	4	ABL20644	Drosophila
413	37	50.0	9325	14	ACL64440	M. xanthu
414	37	50.0	9889	4	ABL10070	Drosophila
415	37	50.0	11739	2	AAK13087	Enterococ
416	37	50.0	11739	6	ABX98882	Enterococ
417	37	50.0	12413	4	ABL20646	Drosophila
418	37	50.0	12619	5	AAH26493	Human low
419	37	50.0	15140	14	ABE11265	Bovine RS
420	37	50.0	15413	4	AAK84002	Human imm
421	37	50.0	21844	5	ABA20799	Human ner
422	37	50.0	22573	2	AAK37254	Human gen
423	37	50.0	22671	2	AAV31200	E. coli J
424	37	50.0	28614	11	ACN44532	Mouse gen
425	37	50.0	40433	6	ABN96830	Gene #332
426	37	50.0	42034	12	ADQ97373	Human can
427	37	50.0	52872	4	ABL04952	Drosophila
428	37	50.0	57763	9	ADA02528	Human MYB
429	37	50.0	57763	10	ADB72266	Human MYB
430	37	50.0	57763	10	ADG95776	Human MYB
431	37	50.0	90220	6	ABK33576	Human cdn
432	37	50.0	110000	2	AAK20248	Continuation (8 of
433	37	50.0	110000	2	AAZ01425	Complete
434	37	50.0	110000	11	ACN43984	Continuation (4 of
435	37	50.0	110000	12	ADN46845	Continuation (19 o
436	37	50.0	110000	12	ADN47591	Continuation (2 of
437	37	50.0	110000	12	ADN46123	Continuation (19 o
438	37	50.0	110000	12	ADN47209	Continuation (2 of
439	37	50.0	110000	12	ADN46464	Continuation (19 o
440	37	50.0	110000	12	ADN47960	Continuation (2 of
441	37	50.0	168174	6	ABT11173	Human 5-1
442	37	50.0	168273	6	ABT11114	Human 5-1

ALIGNMENTS

RESULT 1	
AAF74305	
ID	AAF74305 standard; DNA; 252 BP.
XX	
AC	AAF74305;
XX	
DT	04-MAY-2001 (first entry)
XX	
DE	Canine interleukin-5 coding sequence #2.
XX	
KW	Dog; interleukin-5; IL-5; allergy; cancer; gene therapy;

KW inflammatory reaction; ds.
 XX
 OS Canis sp.
 PN WO200111049-A2.
 XX
 PD 15-FEB-2001.
 XX
 XX 09-AUG-2000; 2000WO-US021651.
 XX
 PF 10-AUG-1999; 99US-00371615.
 XX
 XX (IDEX-) IDEXX LAB INC.
 XX
 PI Guo H, Lawton R, Mermer B, Aiyappa AP;
 XX
 XX WPI; 2001-191542/19.
 DR P-PSDB; AAB72616.
 XX
 XX Novel canine interleukin 5 polynucleotide and polypeptides are used for
 PT generating antibodies which are useful in treating allergies in dogs.
 XX
 PS Example 1; Fig 1; 48pp; English.
 XX
 CC The present invention provides the protein and coding sequences of the
 CC canine interleukin-5 (IL-5) protein. This can be used to treat allergies,
 CC cancer and inflammatory reactions in dogs. The present sequence is one
 CC version of the IL-5 coding sequence shown in the specification
 XX
 SQ Sequence 252 BP; 69 A; 54 C; 60 G; 69 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 6.37e-06 Length: 252
 Score: 74.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0
 US-10-787-382-20 (1-15) x AAF74305 (1-252)
 QY 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
 DB 58 TTTGCTGTAGAAATCCCATGATAGACTGTGGCAGACCTTG 102
 RESULT 2
 AAZ55550
 ID AAZ55550 standard; cDNA; 345 BP.
 XX
 AC AAZ55550;
 XX
 XX 14-MAR-2000 (first entry)
 XX
 DE Canine mature interleukin-5 (IL-5) cDNA.
 XX
 KW Interleukin-5; IL-5; antibody; canine; inhibitor; immune response;
 KW immunoregulation; tumour; cancer; autoimmune disease; vaccine; ss.
 XX
 OS Canis familiaris.
 XX
 PN WO9961618-A2.
 XX
 PD 02-DEC-1999.
 XX
 XX 28-MAY-1999; 99WO-US011942.
 XX
 XX 29-MAY-1998; 98US-0087306P.
 XX
 PA (HESK-) HESKA CORP.
 XX
 PI Sim G, Yang S, Dreitz MJ, Wonderling RS;
 XX
 XX WPI; 2000-072623/06.
 DR
 XX

DR P-PSDB; AAY58220.
 XX
 PT Nucleic acids encoding immunoregulatory proteins from cats or dogs,
 PT useful for treating or preventing e.g. tumors or autoimmune disease.
 XX
 XX Claim 1h; Page 226-227; 264pp; English.
 PS
 XX Sequences AAZ55546-Z55551 represent cDNA sequences encoding canine
 CC interleukin-5 (IL-5). The invention relates to canine IL-4, canine or
 CC feline Flt-3 ligand, canine or feline CD40, canine or feline CD154 (CD40
 CC ligand), canine IL-5, canine IL-13, feline interferon-alpha (IFN-alpha)
 CC and feline granulocyte macrophage colony-stimulating factor (GM-CSF), and
 CC nucleotides which encode these immunoregulatory proteins. The proteins,
 CC their associated nucleic acids, specific antibodies and inhibitors may be
 CC used as vaccines for therapeutic or prophylactic regulation of an immune
 CC response in animals (particularly cats, dogs, horses and humans). They
 CC may be used to treat autoimmune or infectious diseases including
 CC allergies, tumours, inflammation and graft rejection, and to increase the
 CC response from a co-administered antigen. The nucleotide sequences can
 CC also be used for the recombinant production of a protein, while
 CC nucleotide fragments are useful as probes, as amplification primers and
 CC as sources of inhibitory therapeutics (e.g., antisense oligonucleotides).
 CC The proteins may be used to raise antibodies and to screen for modulators
 CC of activity, while the antibodies may be used in detection, and in drug
 CC targeting
 XX
 SQ Sequence 345 BP; 120 A; 68 C; 78 G; 79 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 9.38e-06 Length: 345
 Score: 74.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 3 Gaps: 0
 US-10-787-382-20 (1-15) x AAZ55550 (1-345)
 QY 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
 DB 1 TTTGCTGTAGAAATCCCATGATAGACTGTGGCAGACCTTG 45
 RESULT 3
 AAZ55551/c
 ID AAZ55551 standard; cDNA; 345 BP.
 XX
 AC AAZ55551;
 XX
 XX 14-MAR-2000 (first entry)
 DT
 XX
 DE Canine mature interleukin-5 (IL-5) cDNA complement.
 XX
 KW Interleukin-5; IL-5; antibody; canine; inhibitor; immune response;
 KW immunoregulation; tumour; cancer; autoimmune disease; vaccine; ss.
 XX
 OS Canis familiaris.
 XX
 PN WO9961618-A2.
 XX
 PD 02-DEC-1999.
 XX
 XX 28-MAY-1999; 99WO-US011942.
 XX
 XX 29-MAY-1998; 98US-0087306P.
 XX
 XX (HESK-) HESKA CORP.
 XX
 PI Sim G, Yang S, Dreitz MJ, Wonderling RS;
 XX
 XX WPI; 2000-072623/06.
 DR P-PSDB; AAY58220.
 XX
 XX Nucleic acids encoding immunoregulatory proteins from cats or dogs,
 PT

PT useful for treating or preventing e.g. tumors or autoimmune disease.
 XX
 XX Claim 1h; Page 228; 264pp; English.
 XX
 XX Sequences AAZ55546-Z55551 represent cDNA sequences encoding canine
 CC interleukin-5 (IL-5). The invention relates to canine IL-4, canine or
 CC feline Flt-3 ligand, canine or feline CD40, canine or feline CD154 (CD40
 CC ligand), canine IL-5, canine IL-13, feline interferon-alpha (IFN-alpha)
 CC and feline granulocyte macrophage colony-stimulating factor (GM-CSF), and
 CC nucleotides which encode these immunoregulatory proteins. The proteins,
 CC their associated nucleic acids, specific antibodies and inhibitors may be
 CC used as vaccines for therapeutic or prophylactic regulation of an immune
 CC response in animals (particularly cats, dogs, horses and humans). They
 CC may be used to treat autoimmune or infectious diseases including
 CC allergies, tumours, inflammation and graft rejection, and to increase the
 CC response from a co-administered antigen. The nucleotide sequences can
 CC also be used for the recombinant production of a protein, while
 CC nucleotide fragments are useful as probes, as amplification primers and
 CC as sources of inhibitory therapeutics (e.g., antisense oligonucleotides).
 CC The proteins may be used to raise antibodies and to screen for modulators
 CC of activity, while the antibodies may be used in detection, and in drug
 CC targeting
 XX
 SQ Sequence 345 BP; 79 A; 78 C; 68 G; 120 T; 0 U; 0 Other;
 XX
 Alignment Scores:
 Pred. No.: 9.38e-06 Length: 345
 Score: 74.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 3 Gaps: 0
 US-10-787-382-20 (1-15) x AAZ55551 (1-345)
 QY 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
 Db 345 TTTCGTGTAGAAATCCCATGAATAGACTGGTGGCAGACCTTG 301
 RESULT 4
 AAZ55548
 ID AAZ55548 standard; cDNA; 402 BP.
 XX
 AC AAZ55548;
 XX
 DT 14-MAR-2000 (first entry)
 XX
 DE Canine interleukin-5 (IL-5) cDNA coding region.
 XX
 KW Interleukin-5; IL-5; antibody; canine; inhibitor; immune response;
 KW immunoregulation; tumour; cancer; autoimmune disease; vaccine; ss.
 XX
 OS Canis familiaris.
 XX
 PN WO9961618-A2.
 XX
 PD 02-DEC-1999.
 XX
 PF 28-MAY-1999; 99WO-US011942.
 XX
 PR 29-MAY-1998; 98US-0087306P.
 XX
 PA (HESK-) HESKA CORP.
 XX
 PI Sim G, Yang S, Dreitz MJ, Wonderling RS;
 XX
 DR WPI; 2000-072623/06.
 DR P-PSDB; AAY58219.
 XX
 XX Nucleic acids encoding immunoregulatory proteins from cats or dogs,
 PT useful for treating or preventing e.g. tumors or autoimmune disease.
 XX
 XX Claim 1h; Page 225; 264pp; English.

XX Sequences AAZ55546-Z55551 represent cDNA sequences encoding canine
 CC interleukin-5 (IL-5). The invention relates to canine IL-4, canine or
 CC feline Flt-3 ligand, canine or feline CD40, canine or feline CD154 (CD40
 CC ligand), canine IL-5, canine IL-13, feline interferon-alpha (IFN-alpha)
 CC and feline granulocyte macrophage colony-stimulating factor (GM-CSF), and
 CC nucleotides which encode these immunoregulatory proteins. The proteins,
 CC their associated nucleic acids, specific antibodies and inhibitors may be
 CC used as vaccines for therapeutic or prophylactic regulation of an immune
 CC response in animals (particularly cats, dogs, horses and humans). They
 CC may be used to treat autoimmune or infectious diseases including
 CC allergies, tumours, inflammation and graft rejection, and to increase the
 CC response from a co-administered antigen. The nucleotide sequences can
 CC also be used for the recombinant production of a protein, while
 CC nucleotide fragments are useful as probes, as amplification primers and
 CC as sources of inhibitory therapeutics (e.g., antisense oligonucleotides).
 CC The proteins may be used to raise antibodies and to screen for modulators
 CC of activity, while the antibodies may be used in detection, and in drug
 CC targeting
 XX
 SQ Sequence 402 BP; 129 A; 79 C; 93 G; 101 T; 0 U; 0 Other;
 XX
 Alignment Scores:
 Pred. No.: 1.13e-05 Length: 402
 Score: 74.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 3 Gaps: 0
 US-10-787-382-20 (1-15) x AAZ55548 (1-402)
 QY 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
 Db 58 TTTCGTGTAGAAATCCCATGAATAGACTGGTGGCAGACCTTG 102
 RESULT 5
 AAZ55549/c
 ID AAZ55549 standard; cDNA; 402 BP.
 XX
 AC AAZ55549;
 XX
 DT 14-MAR-2000 (first entry)
 XX
 DE Canine interleukin-5 (IL-5) cDNA coding region complement.
 XX
 KW Interleukin-5; IL-5; antibody; canine; inhibitor; immune response;
 KW immunoregulation; tumour; cancer; autoimmune disease; vaccine; ss.
 XX
 OS Canis familiaris.
 XX
 PN WO9961618-A2.
 XX
 PD 02-DEC-1999.
 XX
 PF 28-MAY-1999; 99WO-US011942.
 XX
 PR 29-MAY-1998; 98US-0087306P.
 XX
 PA (HESK-) HESKA CORP.
 XX
 PI Sim G, Yang S, Dreitz MJ, Wonderling RS;
 XX
 DR WPI; 2000-072623/06.
 DR P-PSDB; AAY58219.
 XX
 XX Nucleic acids encoding immunoregulatory proteins from cats or dogs,
 PT useful for treating or preventing e.g. tumors or autoimmune disease.
 XX
 XX Claim 1h; Page 226; 264pp; English.
 XX
 XX Sequences AAZ55546-Z55551 represent cDNA sequences encoding canine
 CC interleukin-5 (IL-5). The invention relates to canine IL-4, canine or

CC feline Flt-3 ligand, canine or feline CD40, canine or feline CD154 (CD40
 CC ligand), canine IL-5, canine IL-13, feline interferon-alpha (IFN-alpha)
 CC and feline granulocyte macrophage colony-stimulating factor (GM-CSF), and
 CC nucleotides which encode these immunoregulatory proteins. The proteins,
 CC their associated nucleic acids, specific antibodies and inhibitors may be
 CC used as vaccines for therapeutic or prophylactic regulation of an immune
 CC response in animals (particularly cats, dogs, horses and humans). They
 CC may be used to treat autoimmune or infectious diseases including
 CC allergies, tumours, inflammation and graft rejection, and to increase the
 CC response from a co-administered antigen. The nucleotide sequences can
 CC also be used for the recombinant production of a protein, while
 CC nucleotide fragments are useful as probes, as amplification primers and
 CC as sources of inhibitory therapeutics (e.g., antisense oligonucleotides).
 CC The proteins may be used to raise antibodies and to screen for modulators
 CC of activity, while the antibodies may be used in detection, and in drug
 CC targetting

SQ Sequence 402 BP; 101 A; 93 C; 79 G; 129 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 1.13e-05 Length: 402
 Score: 74.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 3 Gaps: 0

US-10-787-382-20 (1-15) x AA255549 (1-402)

Qy 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
 Db 345 TTTGCTGTAGAAATCCCATGATAGACTGTGGCAGACCTTG 301

RESULT 6
 AA255549
 ID AA255549 standard; DNA; 405 BP.

AC AA255549;

DT 04-MAY-2001 (first entry)

DE Canine interleukin-5 coding sequence #1.

DE Dog; interleukin-5; IL-5; allergy; cancer; gene therapy;
 inflammatory reaction; ds.

OS Canis sp.

PN WO20011049-A2.

PD 15-FEB-2001.

PF 09-AUG-2000; 2000WO-US021651.

PR 10-AUG-1999; 98US-00371615.

PA (INDEX-) IDEXX LAB INC.

PI Guo H, Lawton R, Mermer B, Aiyappa AP;

DR WPI; 2001-191542/19.

DR P-PSDB; AAB72615.

PT Novel canine interleukin 5 polynucleotide and polypeptides are used for
 generating antibodies which are useful in treating allergies in dogs.

XX Claim 31; Page 46; 48pp; English.

XX The present invention provides the protein and coding sequences of the
 CC canine interleukin-5 (IL-5) protein. This can be used to treat allergies,
 CC cancer and inflammatory reactions in dogs. The present sequence is one
 CC version of the IL-5 coding sequence shown in the specification

SQ Sequence 405 BP; 131 A; 77 C; 94 G; 103 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.14e-05 Length: 405
 Score: 74.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0

US-10-787-382-20 (1-15) x AAF74300 (1-405)

Qy 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
 Db 58 TTTGCTGTAGAAATCCCATGATAGACTGTGGCAGACCTTG 102

RESULT 7

AA255546

ID AA255546 standard; cDNA; 610 BP.

AC AA255546;

DT 14-MAR-2000 (first entry)

DE Canine interleukin-5 (IL-5) cDNA.

DE Interleukin-5; IL-5; antibody; canine; inhibitor; immune response;
 immunoregulation; tumour; cancer; autoimmune disease; vaccine; ss.

OS Canis familiaris.

PH Key Location/Qualifiers
 FT CDS 29..433
 FT /*tag= a
 FT /product= "Canine IL-5"

PN WO9961618-A2.

PD 02-DEC-1999.

PF 28-MAY-1999; 99WO-US011942.

PR 29-MAY-1998; 98US-0087306P.

PA (HESK-) HESKA CORP.

PI Sim G, Yang S, Dreitz MJ, Wonderling RS;

DR WPI; 2000-072623/06.

DR P-PSDB; AAY58219.

PT Nucleic acids encoding immunoregulatory proteins from cats or dogs,
 useful for treating or preventing e.g. tumors or autoimmune disease.

PS Claim 1h; Page 223-224; 264pp; English.

XX Sequences AA255546-255551 represent cDNA sequences encoding canine
 CC interleukin-5 (IL-5). The invention relates to canine IL-4, canine or
 CC feline Flt-3 ligand, canine or feline CD40, canine or feline CD154 (CD40
 CC ligand), canine IL-5, canine IL-13, feline interferon-alpha (IFN-alpha)
 CC and feline granulocyte macrophage colony-stimulating factor (GM-CSF), and
 CC nucleotides which encode these immunoregulatory proteins. The proteins,
 CC their associated nucleic acids, specific antibodies and inhibitors may be
 CC used as vaccines for therapeutic or prophylactic regulation of an immune
 CC response in animals (particularly cats, dogs, horses and humans). They
 CC may be used to treat autoimmune or infectious diseases including
 CC allergies, tumours, inflammation and graft rejection, and to increase the
 CC response from a co-administered antigen. The nucleotide sequences can
 CC also be used for the recombinant production of a protein, while
 CC nucleotide fragments are useful as probes, as amplification primers and
 CC as sources of inhibitory therapeutics (e.g., antisense oligonucleotides).
 CC The proteins may be used to raise antibodies and to screen for modulators
 CC of activity, while the antibodies may be used in detection, and in drug

CC targetting
 XX Sequence 610 BP; 202 A; 114 C; 139 G; 155 T; 0 U; 0 Other;
 SQ Alignment Scores:
 Pred. No.: 1.89e-05 Length: 610
 Score: 74.00 Matches: 15
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 3 Gaps: 0

US-10-787-382-20 (1-15) x AA255546 (1-610)

QY 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
 DB 86 TTTGCTGTAGAAATCCCATGAATAGACTGGTGGCAGACCTTG 130

RESULT 8
 AA255547/c
 ID AA255547 standard; cDNA; 610 BP.
 XX AC AA255547;
 XX 14-MAR-2000 (first entry)
 XX Canine interleukin-5 (IL-5) cDNA complement.
 XX Interleukin-5; IL-5; antibody; canine; inhibitor; immune response;
 KW immunoregulation; tumour; cancer; autoimmune disease; vaccine; ss.
 XX Canis familiaris.
 XX Key Location/Qualifiers
 FT CDS complement(178..582)
 FT /*tag= a
 FT /product= "Canine IL-5"
 XX WO961618-A2.
 XX 02-DEC-1999.
 XX 28-MAY-1999; 99WO-US011942.
 XX 29-MAY-1998; 98US-0087306P.
 XX (HESK-) HESKA CORP.
 XX Sim G, Yang S, Dretz MJ, Wonderling RS;
 XX WPI; 2000-072623/06.
 XX P-PSDB; AAY58219.
 XX Nucleic acids encoding immunoregulatory proteins from cats or dogs,
 XX useful for treating or preventing e.g. tumors or autoimmune disease.
 XX Claim 1h; Page 224-225; 264pp; English.

Sequences AA255546-Z55551 represent cDNA sequences encoding canine interleukin-5 (IL-5). The invention relates to canine IL-4, canine or feline Flt-3 ligand, canine or feline CD40, canine or feline CD154 (CD40 ligand), canine IL-5, canine IL-13, feline interferon-alpha (IFN-alpha) and feline granulocyte macrophage colony-stimulating factor (GM-CSF), and nucleotides which encode these immunoregulatory proteins. The proteins, their associated nucleic acids, specific antibodies and inhibitors may be used as vaccines for therapeutic or prophylactic regulation of an immune response in animals (particularly cats, dogs, horses and humans). They may be used to treat autoimmune or infectious diseases including allergies, tumours, inflammation and graft rejection, and to increase the response from a co-administered antigen. The nucleotide sequences can also be used for the recombinant production of a protein, while nucleotide fragments are useful as probes, as amplification primers and as sources of inhibitory therapeutics (e.g., antisense oligonucleotides).

CC The proteins may be used to raise antibodies and to screen for modulators
 CC of activity, while the antibodies may be used in detection, and in drug
 CC targetting
 XX Sequence 610 BP; 155 A; 139 C; 114 G; 202 T; 0 U; 0 Other;
 SQ Alignment Scores:
 Pred. No.: 1.89e-05 Length: 610
 Score: 74.00 Matches: 15
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 3 Gaps: 0

US-10-787-382-20 (1-15) x AA255547 (1-610)

QY 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
 DB 525 TTTGCTGTAGAAATCCCATGAATAGACTGGTGGCAGACCTTG 481

RESULT 9
 AA244265
 ID AA244265 standard; DNA; 838 BP.
 XX AC AA244265;
 XX 31-MAR-2000 (first entry)
 XX Porcine IL-5 DNA.
 XX Pig; vaccine; cysticercosis; protective antigen; CC1; CC3; CC4;
 KW tenial cysticercus; gamma interferon; IFN-gamma; interleukin 5; IL-5; ss.
 XX Sus scrofa.
 XX CN1231339-A.
 XX 13-OCT-1999.
 XX 29-JAN-1999; 99CN-00113447.
 XX 29-JAN-1999; 99CN-00113447.
 XX (UYTW-) UNIV NO 2 MILITARY MEDICAL PLA.
 XX Sun S, Dai J;
 XX WPI; 2000-087904/08.
 XX Nucleic acid vaccine for cysticercosis co-contracted by human and pig.
 XX Claim 3; Page 9; 21pp; Chinese.

This invention describes a novel nucleic acid vaccine for preventing and curing human and pork cysticercosis. The invention involves the formation of a eukaryotic expression plasmid from fusion transcript expression unit consisting of three protective antigen genes (CC1, CC3 and CC4) of pig tenial cysticercus and coexpression unit of related cell factor gamma interferon (IFN-gamma) and pork interleukin 5 (IL-5) genes. The CC production and purification process of said nucleic acid vaccine is simple and convenient, the physical and chemical properties of the vaccine are stable, and the vaccine is easy to store and transport, and possesses effective immunological protective function for human and pig cysticercosis. This sequence represents the pig IL-5 gene used in the method of the invention

Sequence 838 BP; 280 A; 148 C; 171 G; 239 T; 0 U; 0 Other;
 SQ Alignment Scores:
 Pred. No.: 0.0211 Length: 838
 Score: 60.00 Matches: 12
 Percent Similarity: 100.00% Conservatives: 2
 Best Local Similarity: 85.71% Mismatches: 0

Query Match: 81.08% Indels: 0
 DB: 3 Gaps: 0

US-10-787-382-20 (1-15) x AA244265 (1-838)

Qy 2 AlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
 |||||:::|||||
 Db 105 GCTGTACAAAGTCCCATGAATAGGCTGGTGGCAGACCTTG 146

RESULT 10
 AAT50756
 ID AAT50756 standard; cDNA; 399 BP.
 XX
 AC AAT50756;
 XX
 DT 17-OCT-2003 (revised)
 DT 24-SEP-1997 (first entry)
 XX
 XX Ovine IL-5 cDNA.
 DE Cytokine; ovine; sheep; interleukin-5; interleukin-12; IL-5; IL-12;
 KW livestock; cow; stress; transport; vaccine adjuvant; veterinary; cancer;
 KW immunosuppression; allergy; reproductive system; growth; early maturity;
 KW antibody; diagnosis; immunopotentiator;
 KW early haematopoietic progenitor cell; cytotoxic cell; thymocyte;
 KW secretion; IgM; IgA; bacterial endotoxin; gamma-interferon; ss.
 XX
 OS Ovis aries.
 XX
 PN WO9700321-A1.
 XX
 PD 03-JAN-1997.
 XX
 PF 14-JUN-1996; 96WO-AU000360.
 XX
 PR 14-JUN-1995; 95AU-00003502.
 PR 27-OCT-1995; 95AU-00006244.
 XX
 PA (CSIR) COMMONWEALTH SCI & IND RES ORG.
 XX
 PI Seow H, Wood P;
 XX
 DR WPI; 1997-077528/07.
 DR P-PSDB; AAW08479.
 XX
 PT Nucleic acid encoding ovine interleukin-5 or -12 - used as vaccine
 PT adjuvants and to treat or prevent microbial infections in livestock.
 XX
 PS Claim 6; Page 41-42; 78pp; English.
 XX
 CC The sequences given in AAT50755-56 encode ovine interleukin-5 (IL-5).
 CC Ovine IL-5 or IL-12 are used to treat and/or prevent infections in
 CC livestock (esp. cows and sheep), particularly where the animals are
 CC stressed, e.g. during transport. IL-5 and IL-12 can also be used as
 CC adjuvants in vaccines for veterinary use (partic. weakly immunogenic
 CC subunit or synthetic peptide vaccines). They may also be used to treat
 CC cancer, immunosuppression and allergy, to enhance/suppress the
 CC reproductive system and to promote growth or early maturity. Optionally
 CC interleukin can be delivered from constructs or delivery cells and
 CC antibodies are useful in enzyme immunoassays for rapid diagnosis of
 CC infection. The interleukins are immunopotentiators, especially IL-5
 CC promotes growth of early haematopoietic progenitor cells and generation
 CC of cytotoxic cells from thymocytes, also it stimulates production and
 CC secretion of IgM and IgA (in synergism with bacterial endotoxin). IL-12
 CC induces production of gamma-interferon by, and proliferation of, T and NK
 CC cells and increases the (non-)specific cytolytic lymphocyte response. The
 CC genetic constructs can also be used for in vitro production of IL-5 or -
 CC 12. (Updated on 17-OCT-2003 to standardise OS field)

SQ Sequence 399 BP; 130 A; 77 C; 93 G; 99 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 0.0903 Length: 399

Score: 55.00 Matches: 12
 Percent Similarity: 92.86% Conservative: 1
 Best Local Similarity: 85.71% Mismatches: 1
 Query Match: 74.32% Indels: 0
 DB: 2 Gaps: 0

US-10-787-382-20 (1-15) x AAT50756 (1-399)

Qy 2 AlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
 |||||:::|||||
 Db 55 GCTGTAGAAAGTACCATGAATAGACTGGTGGCAGACCTTG 96

RESULT 11
 AAT50755
 ID AAT50755 standard; DNA; 520 BP.
 XX
 AC AAT50755;
 XX
 DT 17-OCT-2003 (revised)
 DT 24-SEP-1997 (first entry)
 XX
 XX Ovine IL-5 gene.
 DE Cytokine; ovine; sheep; interleukin-5; interleukin-12; IL-5; IL-12;
 KW livestock; cow; stress; transport; vaccine adjuvant; veterinary; cancer;
 KW immunosuppression; allergy; reproductive system; growth; early maturity;
 KW antibody; diagnosis; immunopotentiator;
 KW early haematopoietic progenitor cell; cytotoxic cell; thymocyte;
 KW secretion; IgM; IgA; bacterial endotoxin; gamma-interferon; ss.
 XX
 OS Ovis aries.
 XX
 PN WO9700321-A1.
 XX
 PD 03-JAN-1997.
 XX
 PF 14-JUN-1996; 96WO-AU000360.
 XX
 PR 14-JUN-1995; 95AU-00003502.
 PR 27-OCT-1995; 95AU-00006244.
 XX
 PA (CSIR) COMMONWEALTH SCI & IND RES ORG.
 XX
 PI Seow H, Wood P;
 XX
 DR WPI; 1997-077528/07.
 DR P-PSDB; AAW08479.
 XX
 PT Nucleic acid encoding ovine interleukin-5 or -12 - used as vaccine
 PT adjuvants and to treat or prevent microbial infections in livestock.
 XX
 PS Claim 6; Page 39-40; 78pp; English.
 XX
 CC The sequences given in AAT50755-56 encode ovine interleukin-5 (IL-5).
 CC Ovine IL-5 or IL-12 are used to treat and/or prevent infections in
 CC livestock (esp. cows and sheep), particularly where the animals are

CC stressed, e.g. during transport. IL-5 and IL-12 can also be used as
 CC adjuvants in vaccines for veterinary use (partic. weakly immunogenic
 CC subunit or synthetic peptide vaccines). They may also be used to treat
 CC cancer, immunosuppression and allergy, to enhance/suppress the
 CC reproductive system and to promote growth or early maturity. Optionally
 CC interleukin can be delivered from constructs or delivery cells and
 CC antibodies are useful in enzyme immunoassays for rapid diagnosis of
 CC infection. The interleukins are immunopotentiators, especially IL-5
 CC promotes growth of early haematopoietic progenitor cells and generation
 CC of cytotoxic cells from thymocytes, also it stimulates production and
 CC secretion of IgM and IgA (in synergism with bacterial endotoxin). IL-12
 CC induces production of gamma-interferon by, and proliferation of, T and NK
 CC cells and increases the (non-)specific cytolytic lymphocyte response. The
 CC genetic constructs can also be used for in vitro production of IL-5 or -
 CC 12. (Updated on 17-OCT-2003 to standardise OS field)

XX
 SQ Sequence 520 BP; 166 A; 99 C; 124 G; 131 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 0.125 Length: 520
 Score: 55.00 Matches: 12
 Percent Similarity: 92.86% Conservative: 1
 Best Local Similarity: 85.71% Mismatches: 1
 Query Match: 74.32% Indels: 0
 DB: 2 Gaps: 0

US-10-787-382-20 (1-15) x AAT50755 (1-520)

QY 2 AlaValGluAAsnProMetAenArgLeuValAlaGluThrIeu 15
 DB 100 GCTGTAGAAAGTACCATGACTGCTGGCAGACCTTG 141

RESULT 12

ABX04971
 ID ABX04971 standard; DNA; 103599 BP.

AC ABX04971,

DT 16-JAN-2003 (first entry)

XX S. cinnamomensis monensin type I polyketide synthase gene cluster.

XX Monensin; gene; cluster; polyketide synthase; antibiotic; ds;
 KW antihelminthic; insecticide; immunosuppressant; antifungal;
 KW antibacterial; polyether; mon BI; mon BII; mon CI; mon CII; mon H;
 KW mon RI; mon RII; mon T; mon AIX; mon AX.

XX Streptomyces cinnamomensis.

XX W0200168867-A1.

XX 20-SEP-2001.

XX 30-MAY-2000; 2000MO-GB002072.

XX 28-MAY-1999; 99GB-00012563.

XX (BIOT-) BIOTICA TECHNOLOGY LTD.

XX Leadlay PF, Staunton J, Oliynyk M;

XX WPI; 2001-611393/70.

DR P-PSDB; ABG99854, ABG99855, ABG99856, ABG99857, ABG99858, ABG99859,
 DR ABG99860, ABG99861, ABG99862, ABG99863, ABG99864, ABG99865, ABG99866,
 DR ABG99867, ABG99868, ABG99869, ABG99870, ABG99871, ABG99872, ABG99873,
 DR ABG99874, ABG99875, ABG99876, ABG99877, ABG99878, ABG99879, ABG99880,
 DR ABG99881, ABG99882, ABG99883, ABG99884, ABG99885, ABG99886, ABG99887.

XX New DNA sequence encoding polyketide synthase, useful for the production
 PT of polyketides such as antibiotic monensin.

XX Claim 1; Page 116-195; 212pp; English.

XX

CC The invention relates to a DNA sequence which is a fully defined sequence
 CC of 103551 base pairs appearing as ABX04971, or its variant, that it is
 CC not a sequence encoding all or part amino acids 1-920 encoded by mon AI
 CC as given in the specification. The DNA is the S. cinnamomensis polyketide
 CC antibiotic monensin biosynthetic gene cluster. Also included are a
 CC recombinant cloning or expression vector comprising the gene cluster, a
 CC transformed host cell which has been transformed to contain the gene
 CC cluster (and is capable of expressing a corresponding polypeptide), a
 CC hybridization probe derived from the gene cluster (for identification and
 CC isolation of the same or analogous gene cluster, e.g. one which binds
 CC specifically to a region of the monensin gene cluster selected from mon
 CC BI, mon BII, mon CI, mon CII, mon H, mon RI, mon RII, mon T, mon AIX and
 CC mon AX), the use of the mon RI gene or variant and a monensin promoter to
 CC control expression of a heterologous gene in Streptomyces cinnamomensis,
 CC a polypeptide encoded by a portion of the monensin gene cluster
 CC (preferably comprising mon BI, mon BII, mon AIX or mon AX or their
 CC mutants, alleles or variants), an epoxidase enzyme encoded by mon CI, a
 CC cyclase enzyme encoded by mon CII, producing S. cinnamomensis capable of
 CC enhanced levels of production of monensin comprising engineering it to
 CC overexpress the mon RI gene, S. cinnamomensis containing multiple copies
 CC of the mon RI gene and/or its variants, expressing a gene heterologous to
 CC S. cinnamomensis comprising transforming S. cinnamomensis with DNA
 CC encoding a heterologous gene and expressing the gene under control of the
 CC activator gene mon RI or actII/orf4 and 13-propyl erythromycin A. The
 CC processes and materials (enzyme systems, nucleic acids and vectors) are
 CC useful for preparing polyketides by recombinant synthesis. The
 CC polyketides are useful as insecticides, antibiotics, antihelminthics,
 CC antifungals, antibacterials or other pharmaceuticals. In particular the
 CC gene is useful for the production of monensin, an antibiotic polyether
 CC polyketide. The present sequence represents the monensin gene cluster
 XX
 SQ Sequence 103599 BP; 13980 A; 37023 C; 37799 G; 14795 T; 0 U; 2 Other;

Alignment Scores:
 Pred. No.: 6.01e+03 Length: 103599
 Score: 46.00 Matches: 9
 Percent Similarity: 84.62% Conservative: 2
 Best Local Similarity: 69.23% Mismatches: 2
 Query Match: 62.16% Indels: 0
 DB: 4 Gaps: 0

US-10-787-382-20 (1-15) x ABX04971 (1-103599)

QY 1 PheAlaValGluAAsnProMetAenArgLeuValAlaGlu 13
 DB 98149 TTCTCTGTGCACGAGCCGATGTCACGTCGTGGTCGCCGA 98187

RESULT 13

ADP69744/C

ID ADP69744 standard; DNA; 312477 BP.

AC ADP69744;

DT 09-SEP-2004 (first entry)

XX Human ROCK 1 DNA #2.

XX ds; gene; human; ROCK 1; hyperproliferative disorder; cancer.

XX Homo sapiens.

XX US2004115641-A1.

XX 17-JUN-2004.

XX 11-DEC-2002; 2002US-00317883.

XX 11-DEC-2002; 2002US-00317883.

XX (ISIS-) ISIS PHARM INC.

XX Cowsett LM, Dobie KW;

XX

DR WPI; 2004-449381/42.
XX New oligonucleotide compound that inhibits expression of ROCK 1, useful
PT for preparing a composition for treating hyperproliferative disorder,
PT e.g., cancer.
XX
PS Example 15; SEQ ID NO 11; 195pp; English.
XX
CC The invention relates to a new compound, targeted to a nucleic acid
CC encoding ROCK 1, that specifically hybridizes with the nucleic acid
CC encoding ROCK 1 and inhibits expression of ROCK 1. The oligonucleotide
CC compound is useful for preparing a composition for treating
CC hyperproliferative disorder, e.g. cancer. The present sequence represents
CC human ROCK 1 DNA.
XX
SQ Sequence 312477 BP; 92110 A; 58460 C; 60155 G; 98196 T; 0 U; 3556 Other;

Alignment Scores:
Pred. No.: 2.34e+04 Length: 312477
Score: 46.00 Matches: 7
Percent Similarity: 83.33% Conservative: 3
Best Local Similarity: 58.33% Mismatches: 2
Query Match: 62.16% Indels: 0
DB: 12 Gaps: 0

US-10-787-382-20 (1-15) x ADP69744 (1-312477)
Qy 3 ValGluAenProMetAsnArgLeuValAlaGluThr 14
ID ABA97721 standard; cDNA; 673 BP.
XX ABA97721;
AC ABA97721;
XX
DT 28-JUN-2002 (first entry)
XX
DE Pepper mild mottle virus envelope gene cDNA.
XX
KW Pepper mild mottle virus; PMMV; envelope; Tobamovirus; transgenic plant;
KW plant expression vector; transgenic; gene; ss.
XX
OS Pepper mild mottle virus.
XX
FH Key Location/Qualifiers
FT CDS 1..480
FT /*tag= a
FT /product= "PMMV envelope protein"
XX
XX
PN KR149216-B1.
XX
PD 17-AUG-1998.
XX
PF 11-NOV-1994; 94KR-00029486.
XX
PR 11-NOV-1994; 94KR-00029486.
XX
PA (POHA-) POHANG ENG COLLEGE.
XX
PI Nam H, Choe J, Lee H, Park Y, Kim C;
XX
DR WPI; 2000-252824/22.
DR P-FSDB; ABB08908.
XX
PT ENVELOPE GENE OF PEPPER MILD MOTTLE VIRUS.
XX
PS Claim 1; Page 7; 12pp; Korean.
XX
CC The invention relates to the envelope gene (ABA97721) and protein
CC (ABB08908) from pepper mild mottle virus (PMMV - a member of the
CC Tobamovirus family of single-stranded RNA viruses). The invention also
CC encompasses plant expression vectors which comprise the PMMV envelope
CC gene. The present sequence represents cDNA encoding the PMMV envelope
CC protein
XX
SQ Sequence 673 BP; 178 A; 141 C; 173 G; 181 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 50.3 Length: 673
Score: 43.00 Matches: 9
Percent Similarity: 69.23% Conservative: 0
Best Local Similarity: 69.23% Mismatches: 4
Query Match: 58.11% Indels: 0
DB: 3 Gaps: 0

US-10-787-382-20 (1-15) x ABA97721 (1-673)
Qy 3 ValGluAenProMetAsnArgLeuValAlaGluThrLeu 15
ID ABA97721 standard; cDNA; 673 BP.
XX ABA97721;
AC ABA97721;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human gene NM_005006, SEQ ID NO 19.
XX
KW Human; ds; gene; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
OS Homo sapiens.
XX
FN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
PA (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
DR WPI; 2003-268312/26.
DR GENBANK; NM_005006.
XX
PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
PS Claim 1; Page; 1017pp; English.
XX
CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a

method for identifying a compound or small molecule that regulates the activity in an animal of one or more of the polypeptides given in the specification, a method for identifying a compound useful in treating pain and a pharmaceutical composition comprising the one or more polypeptides or their antibodies. The polynucleotide or the compound that modulates its activity is useful for preparing a medicament for treating pain (e.g. spinal segmental nerve injury (Chung), chronic constriction injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene therapy). The sequence presented is a human DNA (shown in Table 2 of the specification) which encodes one of the polypeptides of the invention which is differentially expressed during pain. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic form directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 2256 BP; 660 A; 429 C; 557 G; 610 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 223 Length: 2256
Score: 43.00 Matches: 8
Percent Similarity: 66.67% Conservative: 2
Best Local Similarity: 53.33% Mismatches: 5
Query Match: 58.11% Indels: 0
DB: 10 Gaps: 0

US-10-787-382-20 (1-15) x ADE54216 (1-2256)

QY 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrIeu 15
Db 1429 TTCATAATTCCAAATCCAGTAACACACTAGTCACCGAATCTTTG 1385

RESULT 16
AAAF67150/c
ID AAF67150 standard; cDNA; 363 BP.
XX
AC AAF67150;
DT 09-APR-2001 (first entry)
XX
DE Novel human polynucleotide, SEQ ID NO: 2906.
XX Human; cytostatic; gene therapy; colon cancer; prostate cancer;
KW breast cancer; lung cancer; cancer detection; ss.
XX Homo sapiens.
OS
XX WO200102568-A2.
PN
XX 11-JAN-2001.
PD
XX 30-JUN-2000; 2000WO-US018374.
PF
XX 02-JUL-1999; 99US-0142310P.
PR 02-JUL-1999; 99US-0142311P.
XX

(CHIR) CHIRON CORP.
(HYSE-) HYSEQ INC.
XX Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J;
PI Kassam A, Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G;
PI Drmanac R, Crkenjakov R, Drmanac S, Dickson M, Labat I;
PI Leshkowitz D, Kita D, Garcia V, Jones LW, Strache-Crain B;
XX
DR WPI; 2001-091805/10.
XX

Library of polynucleotides for diagnosing a cancerous state of a mammalian cell and detecting cancer, particularly of the colon or prostate, comprises 3351 human polynucleotide sequences.

Claim 9; Page 977; 1046pp; English.

The present sequence is one of 3351 sequences in a library of human polynucleotides. The library is used to detect differentially expressed

genes correlated with a cancerous state of a mammalian cell and can detect colon, prostate, breast and lung cancer. The library can be used to produce probes for detection of mRNA and to produce additional copies of the polynucleotides. The probes can be used for chromosome mapping of the polynucleotide and for detection of transcription levels. Ribozymes or antisense oligonucleotides can be generated. The polynucleotides and their gene products are used as genetic or biochemical markers (e.g. in blood or tissues) that will detect the earliest changes along the carcinogenesis pathway and/or monitor the efficacy of therapies and preventive interventions. The polynucleotides, polypeptides and antibodies against them can be used in pharmaceutical compositions to treat the cancers and proliferative disorders such as neoplasia, dysplasia and hyperplasia

Sequence 363 BP; 98 A; 62 C; 84 G; 118 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 37.8 Length: 363
Score: 42.00 Matches: 8
Percent Similarity: 83.33% Conservative: 2
Best Local Similarity: 66.67% Mismatches: 2
Query Match: 56.76% Indels: 0
DB: 5 Gaps: 0

US-10-787-382-20 (1-15) x AAF67150 (1-363)

QY 1 PheAlaValGluAsnProMetAsnArgLeuValAla 12
Db 44 TTTCGAATTGAGATCCCTTCGTCTCTTCGCA 9

RESULT 17
ABQ91289/c
ID ABQ91289 standard; DNA; 507 BP.
XX
AC ABQ91289;
DT 01-OCT-2002 (first entry)
XX
DE M. capsulatus gene #1274 for DNA array.
XX Micro array; gene; ds; differential expression; gene expression.
KW Methylococcus capsulatus.
OS
XX WO200255655-A2.
PN
XX 18-JUL-2002.
PD
XX 14-JAN-2002; 2002WO-NO000019.
PF
XX 12-JAN-2001; 2001NO-00000235.
PR 12-JAN-2001; 2001NO-00000239.
XX
XX (UNIF-) UNIFOB STIFTELSEN UNIV BERGEN.
XX (TIGR-) TIGR.

Birkeland NK, Eidhammer I, Jonassen I, Jensen HB, Lien T;
PI Lillehaug JR, Lossius I, Eisen JA, Fraser CM, Durkin AS;
PI Salzberg SL;
XX WPI; 2002-557818/59.
XX

Novel DNA array useful for determining differential expression of Methylococcus capsulatus genes, comprises polynucleotides or oligonucleotides representative for a selective number of Methylococcus capsulatus genes.

Claim 14; Page 519; 678pp; English.

The invention relates to a novel DNA array giving a representation of a number of Methylococcus capsulatus genes. The method of the invention is useful for determination of the differential expression of the genes of M. capsulatus, and for studying gene expression on a genomic scale and in

The novel method is useful for upregulating the protein activity. The novel method is useful for lowering the rate of IGF1 signaling and therefore useful for suppressing adhesion, invasion and metastasis where BHD is overexpressed, and therefore endocytosis. The method is also useful for elongating IGF1 effects and therefore for increasing bone formation in osteoporosis conditions. This polynucleotide sequence represents the genomic DNA of a mouse EHD1 gene of the invention


```
PF 07-FEB-2002; 2002WO-IB003040.
XX
PR 07-FEB-2001; 2001FR-00001659.
XX
XX (INSP ) INST PASTEUR.
PA (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI Duchaud E, Tsourit S, Glaser P, Frangeul L, Kunst F, Danchin A;
PI Buchrieser C;
XX
XX WPI; 2003-148459/14.
DR
XX
XX Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.
PT
XX
XX Claim 2; SEQ ID NO 6556; 1205pp; French.
PS
XX
XX The invention relates to the isolation of genes and their encoded
CC proteins from Photorhabdus luminescens. The isolated sequences are
CC sources of probes and primers for detecting the genome of P. luminescens
CC and related species; to study polymorphisms; for gene analysis and for
CC detection/amplification of the genes. Antibodies (Ab) raised against the
CC polypeptides encoded by the genes are used for detection/identification
CC of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
CC carry a gene-containing vector are used to select compounds that
CC modulate, regulate, induce or inhibit expression of the genes in plants,
CC animals or microorganisms other than P. luminescens and are able to alter
CC response or sensitivity to toxins and antibiotics produced by P.
CC luminescens. Cells transformed to express the genes are useful for
CC recombinant production of the proteins, particularly toxins and
CC antibacterials useful as insecticides, bactericides and fungicides. The
CC genes, proteins, vectors containing the genes and Ab are also useful
CC therapeutically (to treat microbial infection by bacteria or fungi that
CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as
CC biopesticides. Other uses of the genes and the proteins are as virulence
CC factors and for identifying targets of human diseases for which P.
CC luminescens is a model (particularly plague and whooping cough). This
CC sequence represents one of the isolated P. luminescens genes
XX
SQ Sequence 381 BP; 105 A; 75 C; 96 G; 105 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 64.4 Length: 381
Score: 41.00 Matches: 8
Percent Similarity: 66.67% Conservative: 2
Best Local Similarity: 53.33% Mismatches: 5
Query Match: 55.41% Indels: 0
DB: 10 Gaps: 0

US-10-787-382-20 (1-15) x ACP68089 (1-381)
QY 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
AD007799
DB 262 TTTCGATAGTAAGTAATCCACTGACCGCGGAGTCTCTTAACCTTG 306

RESULT 23
AD007799
ID ADO07799 standard; cDNA; 924 BP.
XX
XX ADO07799;
AC
XX
XX 01-JUL-2004 (first entry)
DT
XX
XX Fly polynucleotide #58.
DE
XX
XX Fly; gene; ss; fat cell number; fat cell size; obesity; diabetes;
KW anorectic; antidiabetic.
KW
XX
XX Diptera.
OS
XX
XX US2004071700-A1.
PN
XX
XX 15-APR-2004.
PD

XX 09-OCT-2002; 2002US-00267502.
PF
XX
XX 09-OCT-2002; 2002US-00267502.
PR
XX
XX (LIFE-) LIFE SCI DEV CORP.
PA
XX
XX Kim J, Galant R;
PI
XX
XX WPI; 2004-328526/30.
DR
XX
XX P-PSDB; ADO08016.
DR
XX
XX Identifying compounds that influence fat cell number or size for treating
PT or preventing obesity or diabetes by exposing the cell to the agent and
PT identifying fat cell number or size relative to cells not exposed to the
PT agent.
XX
XX Claim 1; SEQ ID NO 125; 275pp; English.
PS
XX
XX The invention relates to a method of identifying compounds that influence
CC fat cell number or size comprising providing a cell that expresses a gene
CC and an agent, exposing the cell to the agent and identifying fat cell
CC number or size relative to cells not exposed to the agent. The method
CC also comprises providing an expression vector and an agent, exposing the
CC vector to the agent, detecting a change in expression of the gene
CC relative to expression of the gene in an expression vector not exposed to
CC the agent, treating a subject with the agent and identifying fat cell
CC number or size in the subject. The agent comprises an antisense
CC oligonucleotide. The subject comprises a mammal, preferably a human. The
CC method also comprises providing a polypeptide and an agent, exposing the
CC polypeptide to the agent, detecting binding of the agent to the
CC polypeptide or a change in an activity of the polypeptide, treating a
CC subject with the agent and identifying fat cell number or size in the
CC subject. The agent comprises an antibody. A method of regulating fat cell
CC number or size comprises providing a subject containing fat cells and an
CC agent that changes the expression of a gene, and treating the subject
CC with the agent under conditions so that fat cell size or number in the
CC subject is altered. The method is useful for identifying compounds that
CC influence fat cell number or size, for preparing a composition for
CC treating or preventing obesity or diabetes. This sequence represents fly
CC cDNA used in the scope of the invention.
XX
SQ Sequence 924 BP; 232 A; 254 C; 269 G; 169 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 192 Length: 924
Score: 41.00 Matches: 9
Percent Similarity: 64.29% Conservative: 0
Best Local Similarity: 64.29% Mismatches: 5
Query Match: 55.41% Indels: 0
DB: 12 Gaps: 0

US-10-787-382-20 (1-15) x ADO07799 (1-924)
QY 2 AlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
AD007799
DB 175 GCTGTGAGAAATCCCGACCGGACTCACTCTGGAACACTT 216

RESULT 24
ABL30155
ID ABL30155 standard; DNA; 941 BP.
XX
XX ABL30155;
AC
XX
XX 26-MAR-2002 (first entry)
DT
XX
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 41938.
DE
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
KW
XX
XX Drosophila melanogaster.
OS
XX
```

```

PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
XX PF
XX 23-MAR-2000; 2000US-0191637P.
PR
XX 11-JUL-2000; 2000US-00614150.
PR
XX (PEXE ) PE CORP NY.
PA
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
XX Claim 1; SEQ ID NO 41938; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 941 BP; 234 A; 259 C; 272 G; 176 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 196 Length: 941
Score: 41.00 Matches: 9
Percent Similarity: 64.29% Conservative: 0
Best Local Similarity: 64.29% Mismatches: 5
Query Match: 55.41% Indels: 0
DB: 4 Gaps: 0
US-10-787-382-20 (1-15) x ABL30155 (1-941)
QY 2 AlaValGluAsnProMetAnArgLeuValAlaGluThrIeu 15
Db 192 GCTGTGAGAAATCCCGACGCGACTCACTGTGGAACACTT 233
RESULT 25
ADT42017
ID ADT42017 standard; cDNA; 1545 BP.
XX
XX ADT42017;
AC
XX 02-DEC-2004 (first entry)
DT
XX
XX Bacterial polynucleotide #16768.
DE
XX Recombinant DNA construct; transformed plant; improved plant property;
XX cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
XX pathogen tolerance; pest tolerance; plant disease resistance;
XX cell cycle pathway modification; plant growth regulator;
XX homologous recombination; seed oil yield; protein yield; carbohydrate;
XX nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
XX bacterial polynucleotide; gene; ss.
XX
XX Bacteria.
OS
XX US2003233675-A1.
PN
XX 19-DEC-2003.
XX
XX 20-FEB-2003; 2003US-00369493.
PF
XX
XX 21-FEB-2002; 2002US-0360039P.
XX
XX (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
PI
XX WPI; 2004-061375/06.
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
XX Claim 1; SEQ ID NO 40455; 122pp; English.
XX
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polynucleotide used in
CC the scope of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
XX Sequence 1545 BP; 220 A; 601 C; 525 G; 199 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 361 Length: 1545
Score: 41.00 Matches: 8
Percent Similarity: 78.57% Conservative: 3
Best Local Similarity: 57.14% Mismatches: 3
Query Match: 55.41% Indels: 0
DB: 13 Gaps: 0
US-10-787-382-20 (1-15) x ADT42017 (1-1545)
QY 2 AlaValGluAsnProMetAnArgLeuValAlaGluThrIeu 15
Db 1153 GCCGAGGACAAATCCGATCAACGCGTGTGCGCCGACCCCTG 1194
RESULT 26
ABZ16494/C
ID ABZ16494 standard; DNA; 1798 BP.
XX
XX ABZ16494;
AC
XX 21-JAN-2003 (first entry)
DT
XX
XX Arabidopsis thaliana stress regulated gene SEQ ID NO 4299.
DE
XX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX
XX Arabidopsis thaliana.
OS

```


CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 2941 BP; 835 A; 629 C; 698 G; 779 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 798 Length: 2941
 Score: 41.00 Matches: 9
 Percent Similarity: 64.29% Conservative: 0
 Best Local Similarity: 64.29% Mismatches: 5
 Query Match: 55.41% Indels: 0
 DB: 4 Gaps: 0

US-10-787-382-20 (1-15) x ABL30154 (1-2941)

Qy 2 AlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
 Db 1192 GCTGTGAGAAATCCCGACCGCGACTCACTGTGGAACACTT 1233

RESULT 31
 ADD93578/C
 ID ADD93578 standard; DNA; 3004 BP.
 XX
 AC ADD93578;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Arabidopsis pathogen-inducible promoter P26.
 XX
 KW Promoter; plant; pathogen resistance; transgenic; crop protection;
 KW disease resistance; ds.
 XX
 OS Arabidopsis thaliana.
 XX
 PN WO2003083042-A2.
 XX
 PD 09-OCT-2003.
 XX
 PF 24-OCT-2002; 2002WO-US034220.
 XX
 PR 24-OCT-2001; 2001US-0335249P.
 XX
 PA (DNAP) DNA PLANT TECHNOLOGY CORP.
 XX
 PI Repetti P., Scofield SR, Century K;
 XX
 DR WPI; 2003-779453/73.
 XX
 PT New isolated nucleic acid molecule useful for enhancing resistance of a
 PT plant to a pathogen (e.g. bacteria, viruses, fungi, nematodes or
 PT insects).
 XX
 PS Claim 7; Page 29-30; 41pp; English.
 XX
 CC The present sequence is that of a pathogen-inducible promoter, denoted
 CC P26, of Arabidopsis thaliana. Pathogen-inducible promoters were
 CC identified in an experiment in which gene expression in wild-type Col-0
 CC Arabidopsis plants was monitored over the course of a Botrytis cinerea
 CC infection. Genes were identified whose expression was up-regulated during
 CC pathogen infection compared to uninfected tissue, and promoter sequences
 CC were isolated by PCR amplification. The promoters can be used to enhance
 CC pathogen resistance in plants, e.g. by being operably linked to a
 CC resistance effector gene.
 XX
 SQ Sequence 3004 BP; 989 A; 582 C; 466 G; 967 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 819 Length: 3004
 Score: 41.00 Matches: 6
 Percent Similarity: 80.00% Conservative: 6
 Best Local Similarity: 40.00% Mismatches: 3

Query Match: 55.41% Indels: 0
 DB: 10 Gaps: 0

US-10-787-382-20 (1-15) x ADD93578 (1-3004)

Qy 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
 Db 2964 TTTAATATAGAAATCCCGACACATTAATTAACAACAGTA 2920

RESULT 32
 ADE06707/C
 ID ADE06707 standard; DNA; 5138 BP.
 XX
 AC ADE06707;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE ORF1 gene from Hepatitis E virus strain JRAL.
 XX
 KW ds; gene; hepatitis E virus; chips; kit; detection.
 XX
 OS Hepatitis E virus; strain JRAL.
 XX
 FH Key Location/Qualifiers
 FT CDS 27..5138
 FT /*tag= a
 XX
 PN WO2003000887-A1.
 XX
 PD 03-JAN-2003.
 XX
 PF 25-JUN-2002; 2002WO-JP006365.
 XX
 PR 25-JUN-2001; 2001JP-00191837.
 XX
 PA (TOKE) TOSHIBA KK.
 XX
 PI Takahashi K, Mishiro S, Oota Y, Hashimoto M, Maekubo H;
 XX
 DR WPI; 2003-239144/23.
 DR P-PSDB; ADE06657.
 XX
 PT Polynucleotide probes and primers originating in Japanese hepatitis E
 PT virus, applicable in chips and kits for detecting hepatitis E virus and
 PT its infection.
 XX
 PS Disclosure; SEQ ID NO 1; 238pp; Japanese.
 XX
 CC The invention relates to novel polynucleotide probes for detecting
 CC hepatitis E virus polynucleotides which contain a sequence of at least 8
 CC consecutive nucleotides of the hepatitis E virus genome. The hepatitis E
 CC virus genomic sequences are derived from various strains of the Japanese
 CC Hepatitis E virus. The probes and primers are applicable in chips and
 CC kits for detecting hepatitis E virus and its infection. This sequence
 CC represents the open reading frame 1 from the Hepatitis E virus strain
 CC JRAL.
 XX
 SQ Sequence 5138 BP; 966 A; 1470 C; 1380 G; 1322 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 1.59e+03 Length: 5138
 Score: 41.00 Matches: 8
 Percent Similarity: 73.33% Conservative: 3
 Best Local Similarity: 53.33% Mismatches: 4
 Query Match: 55.41% Indels: 0
 DB: 10 Gaps: 0

US-10-787-382-20 (1-15) x ADE06707 (1-5138)

Qy 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
 Db 248 TTCATTGTGTATACCCCTCTGAATAGGATGGTTCGAAGACCTC 204

```

RESULT 33
ABL30196
ID ABL30196 standard; DNA; 6496 BP.
XX
AC ABL30196;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 42061.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PP 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
XX
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
WPI; 2001-656860/75.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Claim 1; SEQ ID NO 42061; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 6496 BP; 1733 A; 1444 C; 1490 G; 1829 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2.12e+03 Length: 6496
Score: 41.00 Matches: 9
Percent Similarity: 64.29% Conservative: 0
Best Local Similarity: 64.29% Mismatches: 5
Query Match: 55.41% Indels: 0
DB: 4 Gaps: 0

US-10-787-382-20 (1-15) x ABL30196 (1-6496)
QY 2 AlavalGluAenProMetAenArgLeuValAlaGluThrLeu 15
DB 65 GCTGTGAGAAATCCCGACCGCACTCTCTGTGTGAAACACTT 106

RESULT 34
ADE06697/c
ID ADE06697 standard; DNA; 7230 BP.
XX
AC ADE06697;
XX
DT 29-JAN-2004 (first entry)
XX
DE Hepatitis E virus strain JRAL genomic sequence.
XX

```

```

KW hepatitis E virus; chips; kit; detection; ds.
XX
OS Hepatitis E virus; strain JRAL.
XX
PN WO2003000887-A1.
XX
PD 03-JAN-2003.
XX
PP 25-JUN-2002; 2002WO-JP006365.
XX
PR 25-JUN-2001; 2001JP-00191837.
XX
PA (TOKE ) TOSHIBA KK.
XX
PI Takahashi K, Mishiho S, Oota Y, Hashimoto M, Maekubo H;
XX
WPI; 2003-239144/23.
XX
PT Polynucleotide probes and primers originating in Japanese hepatitis E
PT virus, applicable in chips and kits for detecting hepatitis E virus and
PT its infection.
XX
PS Claim 1; SEQ ID NO 48; 238pp; Japanese.
XX
CC The invention relates to novel polynucleotide probes for detecting
CC hepatitis E virus polynucleotides which contain a sequence of at least 8
CC consecutive nucleotides of the hepatitis E virus genome. The hepatitis E
CC virus genomic sequences are derived from various strains of the Japanese
CC Hepatitis E virus. The probes and primers are applicable in chips and
CC kits for detecting hepatitis E virus and its infection. This sequence
CC represents the genomic sequence from the Hepatitis E virus strain JRAL.
XX
SQ Sequence 7230 BP; 1344 A; 2120 C; 1880 G; 1886 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2.42e+03 Length: 7230
Score: 41.00 Matches: 8
Percent Similarity: 73.33% Conservative: 3
Best Local Similarity: 53.33% Mismatches: 4
Query Match: 55.41% Indels: 0
DB: 10 Gaps: 0

US-10-787-382-20 (1-15) x ADE06697 (1-7230)
QY 1 PheAlaValGluAenProMetAenArgLeuValAlaGluThrLeu 15
DB 248 TTCATTGTGTATACCCCTCTGAATAGGATGGTTCCAAAGCACCTC 204

RESULT 35
ADE06694/c
ID ADE06694 standard; DNA; 7233 BP.
XX
AC ADE06694;
XX
DT 29-JAN-2004 (first entry)
XX
DE Hepatitis E virus strain JMY-Haw genomic sequence.
XX
KW hepatitis E virus; chips; kit; detection; ds.
XX
OS Hepatitis E virus; strain JMY-Haw.
XX
PN WO2003000887-A1.
XX
PD 03-JAN-2003.
XX
PP 25-JUN-2002; 2002WO-JP006365.
XX
PR 25-JUN-2001; 2001JP-00191837.
XX
PA (TOKE ) TOSHIBA KK.
XX
PI Takahashi K, Mishiho S, Oota Y, Hashimoto M, Maekubo H;

```

XX WPI; 2003-239144/23.

XX Polynucleotide probes and primers originating in Japanese hepatitis E

XX virus, applicable in chips and kits for detecting hepatitis E virus and

XX its infection.

XX Claim 1; SEQ ID NO 45; 238pp; Japanese.

XX The invention relates to novel polynucleotide probes for detecting

XX hepatitis E virus polynucleotides which contain a sequence of at least 8

XX consecutive nucleotides of the hepatitis E virus genome. The hepatitis E

XX virus genomic sequences are derived from various strains of the Japanese

XX Hepatitis E virus. The probes and primers are applicable in chips and

XX kits for detecting hepatitis E virus and its infection. This sequence

XX represents the genomic sequence from the Hepatitis E virus strain JMY-

XX Haw.

XX Sequence 7233 BP; 1313 A; 2103 C; 1901 G; 1916 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.42e+03	Length:	7233
Score:	41.00	Matches:	7
Percent Similarity:	73.33%	Conservative:	4
Best Local Similarity:	46.67%	Mismatches:	4
Query Match:	55.41%	Indels:	0
DB:	10	Gaps:	0

US-10-787-382-20 (1-15) x ADR06694 (1-7233)

QY 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15

DB 224 TTCATTATGATATACCCGCTGATAGGATGATTCAGAGCACCTC 180

RESULT 36

ADV41357/c

ID ADV41357 standard; cDNA; 10708 BP.

XX AC ADV41357;

XX DT 10-FEB-2005 (first entry)

XX DE Rat cardiotoxicity marker gene, SEQ ID NO:2223.

XX KW Toxicology screening; drug screening; cardiotoxicity, drug-induced;

XX KW expression profile; gene expression; myocarditis; heart arrhythmia;

XX KW tachycardia; myocardial ischemia; angina; hypertension; hypotension;

XX KW dyspnea; cardiogenic shock; cardiovascular disease; gene; ss.

XX OS Rattus norvegicus.

XX PN WO2004063334-A2.

XX PD 29-JUL-2004.

XX PF 08-JAN-2004; 2004WO-US000240.

XX PR 08-JAN-2003; 2003US-00338044.

XX PA (GENE-) GENE LOGIC INC.

XX PI Mendrick DL, Porter MW, Johnson KR, Higgs B, Castle A;

XX PI Elashoff M;

XX WPI; 2004-561893/54.

XX DR REFSEQ; NM_031046.

XX Predicting a toxic effect of a compound by obtaining a gene expression

XX profile of a tissue or cell sample exposed to the compound and comparing

XX the gene expression profile to a database comprising all of the data

XX given in the specification.

XX Example 1; SEQ ID NO 2223; 520pp; English.

XX The invention relates to a method of predicting at least one toxic effect

XX of a compound by obtaining a gene expression profile of a tissue or cell

XX sample exposed to the compound and comparing the gene expression profile

XX to a database of toxicity prediction genes which are differentially

XX expressed on exposure to known toxins. The tissue or cell sample is

XX preferably derived from heart tissue, and the predicted toxic effect is

XX especially a cardiotoxic effect such as myocarditis, arrhythmias,

XX tachycardia, myocardial ischemia, angina, hypertension, hypotension,

XX dypnea, cardiogenic shock or other specific heart pathologies. The

XX invention is based on the elucidation of global changes in gene

XX expression in tissues or cells exposed to known toxins, particularly

XX cardiotoxins, and the identification of individual genes (toxicity

XX markers) that are differentially expressed on toxin exposure. The

XX invention also relates to methods for predicting the progression of a

XX compound's toxic effects, for predicting the cardiotoxicity of a

XX compound, for identifying an agent that modulates the onset or

XX progression of a toxic response, and for predicting which cellular

XX pathways a particular compound will modulate. The invention further

XX relates to sets of at least two hybridization probes specific for

XX toxicity prediction genes, and solid supports and kits comprising them.

XX The method of the invention is useful in toxicology screening for

XX predicting the toxic effects (especially cardiotoxic effects) of

XX compounds such as pharmaceutical agents or environmental pollutants.

XX Sequences ADV9135-ADV41830 represent cardiotoxicity marker

XX polynucleotides of rat origin whose expression is altered on exposure to

XX at least one cardiotoxin. Note: The sequence data for this patent did not

XX form part of the printed specification, but was obtained in electronic

XX format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 10708 BP; 2918 A; 2618 C; 2797 G; 2375 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	3.92e+03	Length:	10708
Score:	41.00	Matches:	10
Percent Similarity:	78.57%	Conservative:	1
Best Local Similarity:	71.43%	Mismatches:	3
Query Match:	55.41%	Indels:	0
DB:	13	Gaps:	0

US-10-787-382-20 (1-15) x ADV41357 (1-10708)

QY 2 AlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15

DB 9621 GCTGTAAACAACCCCATGTGCGCTTGTGCTGAGAGTTTA 9580

RESULT 37

ACF67367_07

Continuation (8 of 57) of ACF67367 from base 700001 (Photographus luminescens nucleotide

WP Sequence split into 57 fragments LOCUS ACF67367 Accession ACF67367

WP	Fragment Name	Begin	End
WP	ACF67367_00	1	110000
WP	ACF67367_01	100001	210000
WP	ACF67367_02	200001	310000
WP	ACF67367_03	300001	410000
WP	ACF67367_04	400001	510000
WP	ACF67367_05	500001	610000
WP	ACF67367_06	600001	710000
WP	ACF67367_07	700001	810000
WP	ACF67367_08	800001	910000
WP	ACF67367_09	900001	1010000
WP	ACF67367_10	1000001	1110000
WP	ACF67367_11	1100001	1210000
WP	ACF67367_12	1200001	1310000
WP	ACF67367_13	1300001	1410000
WP	ACF67367_14	1400001	1510000
WP	ACF67367_15	1500001	1610000
WP	ACF67367_16	1600001	1710000
WP	ACF67367_17	1700001	1810000
WP	ACF67367_18	1800001	1910000
WP	ACF67367_19	1900001	2010000
WP	ACF67367_20	2000001	2110000
WP	ACF67367_21	2100001	2210000

Single exon nucleic acid probes for analyzing gene expression in human hearts.

CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway;
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX

SQ Sequence 480 BP; 176 A; 84 C; 93 G; 127 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 137 Length: 480
Score: 40.00 Matches: 8
Percent Similarity: 73.33% Conservative: 3
Best Local Similarity: 53.33% Mismatches: 4
Query Match: 54.05% Indels: 0
DB: 8 Gaps: 0

US-10-787-382-20 (1-15) x ACA43050 (1-480)

QY 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
||||||| :|||:|||||
Db 299 TTTCGGGTCTCACACCAATATCTTCTTTGATTGTTGAAACGCTG 255

RESULT 43

ABA31096

ID ABA31096 standard; DNA; 567 BP.

XX AC ABA31096;

XX DT 23-JAN-2002 (first entry)

XX DE Probe #9562 for gene expression analysis in human heart cell sample.

XX DE Human; gene expression; heart; microarray; vascular system; probe;

XX DE cardiovascular disease; hypertension; cardiac arrhythmia;

XX DE congenital heart disease; ss.

XX OS Homo sapiens.

XX PN WO200157274-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US0000666.

XX PR 04-FEB-2000; 2000US-0180312P.

XX PR 26-MAY-2000; 2000US-0207456P.

XX PR 30-JUN-2000; 2000US-00608408.

XX PR 03-AUG-2000; 2000US-00632366.

XX PR 21-SEP-2000; 2000US-0234687P.

XX PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WI MPI; 2001-488899/53.

XX Single exon nucleic acid probes for analyzing gene expression in human
PT hearts.

XX Claim 1; SEQ ID NO 9562; 530pp; English.

XX The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart. The
CC present sequence is one such probe. The probes may be used for
CC predicting, measuring and displaying gene expression in samples derived
CC from the human heart via microarrays. By measuring gene expression, the
CC probes are useful for predicting, diagnosing, grading, staging,
CC monitoring and prognosing diseases of the human heart and vascular system
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC congenital heart disease. Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX

SQ Sequence 567 BP; 138 A; 101 C; 157 G; 164 T; 0 U; 7 Other;

Alignment Scores:

Pred. No.: 169 Length: 567
Score: 40.00 Matches: 8
Percent Similarity: 73.33% Conservative: 3
Best Local Similarity: 53.33% Mismatches: 4
Query Match: 54.05% Indels: 0
DB: 4 Gaps: 0

US-10-787-382-20 (1-15) x ABA31096 (1-567)

QY 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
||||||| :|||:|||||

Db 48 TTCCATGTGGAAAGACCTTTGCTAGGGTGTTCACAGGAACTT 92

RESULT 44

ACH72974

ID ACH72974 standard; DNA; 569 BP.

XX AC ACH72974;

XX DT 29-JUL-2004 (first entry)

XX DE Human genome derived single exon probe #6169.

XX DE Human; probe; ss; gene expression; single exon probe; microarray;

XX DE alternative splicing event; genomic alteration.

XX OS Homo sapiens.

XX PN US2003194704-A1.

XX PD 16-OCT-2003.

XX PF 03-APR-2002; 2002US-00029386.

XX PR 03-APR-2002; 2002US-00029386.

XX PA (PENN/) PENN S G.

XX PA (RANK/) RANK D R.

XX PA (HANZ/) HANZEL D K.

XX PI Penn SG, Rank DR, Hanzel DK;

XX WI MPI; 2004-119264/12.

XX New human genome-derived single exon nucleic acid probes useful for human

XX WO200175067-A2.
 XX 11-OCT-2001.
 XX 30-MAR-2001; 2001WO-US008631.
 XX 31-MAR-2000; 2000US-00540217.
 XX 23-AUG-2000; 2000US-00649167.
 XX (HYSE-) HYSEQ INC.
 XX Drmanac RT, Liu C, Tang YT;
 XX WPI; 2001-639362/73.
 XX P-PSDB; ABG09442.
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX Claim 1; SEQ ID NO 9433; 103pp; English.
 XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX SQ Sequence 2520 BP; 714 A; 636 C; 598 G; 569 T; 0 U; 3 Other;
 Alignment Scores:
 Pred. No.: 1.06e+03 Length: 2520
 Score: 40.00 Matches: 8
 Percent Similarity: 69.23% Conservative: 1
 Best Local Similarity: 61.54% Mismatches: 4
 Query Match: 54.05% Indels: 0
 Gaps: 5
 US-10-787-382-20 (1-15) x AAS73629 (1-2520)
 QY 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGlu 13
 Db 2018 TTTCGCTAGAAAAGCCTTTCAGAGGTGTGGTGGAGGAA 1980
 RESULT 49
 ADM01809/c
 ID ADM01809 standard; cDNA; 2591 BP.
 XX
 XX AC ADM01809;
 XX
 XX DT 20-MAY-2004 (first entry)
 XX
 XX DE Human cDNA of the invention SEQ ID NO:494.
 XX
 XX KW ss; gene; human; gene therapy; diagnostic marker; pharmaceutical.

XX Homo sapiens.
 XX EP1347046-A1.
 XX 24-SEP-2003.
 XX 12-APR-2002; 2002EP-00008400.
 XX 22-MAR-2002; 2002JP-00137785.
 XX (REAS-) RES ASSOC BIOTECHNOLOGY.
 XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
 PI Seki N, Yoshikawa T, Otsuka M, Nagahara K, Masuho Y;
 XX WPI; 2003-723558/69.
 XX P-PSDB; ADM04252.
 XX New polynucleotides and polypeptides are useful in gene therapy, for
 PT developing a diagnostic marker or medicines for regulating their
 PT expression and activity, or as a target of gene therapy.
 XX Claim 1; SEQ ID NO 494; 305pp; English.
 XX The invention relates to a novel human polynucleotide and the encoded
 CC polypeptide. A polynucleotide of the invention may have a use in gene
 CC therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful
 CC as a primer for synthesizing the polynucleotide or as a probe for
 CC detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are
 CC useful in gene therapy, for developing a diagnostic marker or medicines
 CC for regulating their expression and activity, or as a target of gene
 CC therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides
 CC are useful as pharmaceutical agents. The present sequence represents a
 CC cDNA sequence of the invention.
 XX SQ Sequence 2591 BP; 744 A; 477 C; 580 G; 790 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.1e+03 Length: 2591
 Score: 40.00 Matches: 7
 Percent Similarity: 83.33% Conservative: 3
 Best Local Similarity: 58.33% Mismatches: 2
 Query Match: 54.05% Indels: 0
 Gaps: 0
 US-10-787-382-20 (1-15) x ADM01809 (1-2591)
 QY 4 GluAsnProMetAsnArgLeuValAlaGluThrLeu 15
 Db 1621 GAGAACCCATTAAACAGTCTTATTGCCAACTGCTG 1586
 RESULT 50
 ABL14942/c
 ID ABL14942 standard; cDNA; 6692 BP.
 XX
 XX AC ABL14942;
 XX
 XX DT 26-MAR-2002 (first entry)
 XX
 XX DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 39308.
 XX
 XX KW Drosophila; developmental biology; cell signalling; insecticide;
 XX pharmaceutical; gene; ss.
 XX Drosophila melanogaster.
 XX WO200171042-A2.
 XX 27-SEP-2001.
 XX 23-MAR-2001; 2001WO-US009231.

```

XX 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
XX PA (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
XX WPI; 2001-656860/75.
XX P-PSDB; ABE70839.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
PT
XX Claim 1; SEQ ID NO 39308; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABR57737-
CC ABR72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 6692 BP; 1978 A; 1255 C; 1270 G; 2189 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 3.53e+03 Length: 6692
Score: 40.00 Matches: 6
Percent Similarity: 83.33% Conservative: 4
Best Local Similarity: 50.00% Mismatches: 2
Query Match: 54.05% Indels: 0
DB: 4 Gaps: 0
US-10-787-382-20 (1-15) x ABL14942 (1-6692)
QY 3 ValGluAsnProMetAsnArgLeuValAlaGluThr 14
DB 4781 ATCAACAACCCATTGATCGATTGTTGATCTAGACA 4746
Search completed: December 21, 2005, 17:16:44
Job time : 453 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 21, 2005, 17:06:14 ; Search time 2610 Seconds
(without alignments)
268.891 Million cell updates/sec

Title: US-10-787-382-20

Perfect score: 74

Sequence: 1 FAVENPMNRLVAETL 15

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10787382/runat 16122005 132436 8024/app query fasta_1.199
-DB=EST -QFMT=fasCap -SURFIX=p2n.rst -MINMATCH=0.1 -LOOFCU=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=500
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=50 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10787382@cgn 1 1 5315 @runat 16122005 132436 8024 -NCPU=6 -ICPU=3
-NO MMAP -LARGESOURCY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_est3:*

4: gb_hc3:*

5: gb_hc4:*

6: gb_est5:*

7: gb_est6:*

8: gb_est7:*

9: gb_gss1:*

10: gb_gss2:*

11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	ID	Description

C	1	47	63.5	541	4	CNSOABAJ BX817426 Arabidops
C	2	46	62.2	376	1	AI002749 am98a04.8
C	3	46	62.2	565	7	COL92024 EC32618.5
C	4	46	62.2	874	11	CR821531 GR0AA50D
C	5	45	60.8	256	1	AV222732 AV222732
C	6	45	60.8	438	8	DN847029 KECB30-27
C	7	45	60.8	490	7	CN755695 ID0AAA16C

CF587487	USDA-FP 1	504	60.8	45	6	CF587487
CN753056	APHL3LD-X	539	60.8	45	7	CN753056
CN754327	ID0AAA12C	625	60.8	45	7	CN754327
CV834660	ID0ACC26B	856	60.8	45	8	CV834660
CV835366	ID0ACC28B	902	60.8	45	8	CV835366
CG458219	PUPXV93TB	959	60.8	45	10	CG458219
CA58198	EST0003873	1000	60.8	45	6	CA58198
B1084264	602869768	1012	60.8	45	2	B1084264
BP886906	BP886906	471	59.5	44	3	BP886906
BM281788	KI20G03.Y	505	59.5	44	3	BM281788
CO675844	BN40.046B	510	59.5	44	6	CO675844
CG830599	CG830599	572	59.5	44	7	CG830599
CG132831	Pan trogl	675	59.5	44	10	CG132831
B173326	603353349	797	59.5	44	3	B173326
CL670791	PR10163A	817	59.5	44	10	CL670791
CL665378	PR10149C	832	59.5	44	10	CL665378
BF977950	602148422	966	59.5	44	2	BF977950
AL430930	clone XBA	1011	59.5	44	11	CNS0764C
AL347462	Tetraodon	1016	59.5	44	11	CNS05P3H
DN684850	CGX38-D04	1415	58.8	43	8	DN684850
CL291538	ZMMBB0063	684	58.1	43	10	CL291538
CZ491475	OA_BBA014	362	58.1	43	10	CZ491475
CB827405	LjNEST74d	418	58.1	43	6	CB827405
DR897668	JGI_XZ742	494	58.1	43	8	DR897668
CE797827	tigr-988-	511	58.1	43	10	CE797827
CC062239	ugma001e0	517	58.1	43	9	CC062239
CZ331571	ZMMBF0039	522	58.1	43	10	CZ331571
CA877602	KO958A01-	536	58.1	43	6	CA877602
CC578117	CH240.456	553	58.1	43	9	CC578117
CG841296	Ynhw1910	562	58.1	43	10	CG841296
AZ759493	IM0552K03	591	58.1	43	9	AZ759493
CF799042	Lr_PACFC	598	58.1	43	6	CF799042
CG950325	MBEFH66TF	608	58.1	43	10	CG950325
CM171674	104_582.1	613	58.1	43	10	CM171674
CM301651	104_785-1	639	58.1	43	10	CM301651
CM277572	104_751-1	665	58.1	43	10	CM277572
BB303483	BB303483	678	58.1	43	2	BB303483
BH842461	TC3-51G7.	682	58.1	43	9	BH842461
BH845769	TC3-51G7.	682	58.1	43	9	BH845769
CZ422877	1018580.R	683	58.1	43	10	CZ422877
BB344850	BB344850	708	58.1	43	2	BB344850
AG170565	Pan trogl	720	58.1	43	10	AG170565
AG175393	Pan trogl	722	58.1	43	10	AG175393
BI912325	603067613	726	58.1	43	3	BI912325
CC339087	OGPBC03TH	734	58.1	43	9	CC339087
CM277571	104_751.1	742	58.1	43	10	CM277571
CC824344	ZMMBB016	761	58.1	43	9	CC824344
BH924917	odh76G03.	768	58.1	43	9	BH924917
CV266035	WS02028.B	768	58.1	43	7	CV266035
CR839033	GROAAA74B	769	58.1	43	11	CR839033
CL676674	PR10119b	864	58.1	43	10	CL676674
AZ539120	ENTGD12TF	877	58.1	43	9	AZ539120
CG379949	OGABK59TC	899	58.1	43	7	CG379949
BM544082	AGENCOURT	905	58.1	43	10	BM544082
DR740001	FGAS00026	1063	58.1	43	3	DR740001
BO684395	AGENCOURT	1126	58.1	43	8	BO684395
AK047369	Mus muscu	1252	58.1	43	5	AK047369
BC028017	Homo sapi	1411	58.1	43	4	BC028017
AK045349	Mus muscu	1560	58.1	43	4	AK045349
CC127275	NDL.87B20	3400	56.8	42	9	CC127275
AV167716	AV167716	305	56.8	42	1	AV167716
AV296711	AV296711	324	56.8	42	1	AV296711
BB465289	BB465289	326	56.8	42	2	BB465289
BB876839	hr33h05.b	326	56.8	42	9	BB876839
BB391641	BB391641	337	56.8	42	2	BB391641
BY649054	BY649054	393	56.8	42	5	BY649054
BQ078906	EST594313	415	56.8	42	3	BQ078906
BJ454136	BJ454136	425	56.8	42	3	BJ454136
BF042251	BP250004B	472	56.8	42	2	BF042251
CB436836	684000.MA	490	56.8	42	6	CB436836
CN755347	ID0AAA15C	490	56.8	42	7	CN755347
CB355254	ZF001-P00	505	56.8	42	6	CB355254

C 81	42	56.8	542	3	BM343871	BM343871 rr44a01.y	C 154	41	55.4	560	11	CR244142	CR244142 Forward s
C 82	42	56.8	579	3	BM092916	BM092916 saj02f09.	C 155	41	55.4	566	9	BH786250	BH786250 fzm0014f0
C 83	42	56.8	594	3	BM343994	BM343994 rr45d11.y	C 156	41	55.4	568	7	CO102578	CO102578 GR_EB003
C 84	42	56.8	608	10	CW513785	CW513785 115.5.105	C 157	41	55.4	569	9	AZ300131	AZ300131 RPCI-23-1
C 85	42	56.8	622	10	CE738182	CE738182 tigr-g8s-	C 158	41	55.4	573	2	BG520968	BG520968 p803b05.y
C 86	42	56.8	626	10	CE812711	CE812711 tigr-g8s-	C 159	41	55.4	579	9	CC474029	CC474029 CH240.298
C 87	42	56.8	631	1	AW255594	AW255594 ML638 pep	C 160	41	55.4	580	9	CC669780	CC669780 QGVBM48TV
C 88	42	56.8	644	5	BX466398	BX466398 BX466398	C 161	41	55.4	581	1	A1542067	A1542067 SD08134.5
C 89	42	56.8	655	3	BJ277729	BJ277729 BJ277729	C 162	41	55.4	585	10	CW166348	CW166348 104.575.1
C 90	42	56.8	659	9	BH844795	BH844795 TC3-51M10	C 163	41	55.4	586	10	CG845893	CG845893 OeAGC62TV
C 91	42	56.8	664	9	AQ413187	AQ413187 RPCI-11-1	C 164	41	55.4	596	7	CO963847	CO963847 BeE30N10A
C 92	42	56.8	665	6	CA138015	CA138015 SCEPRT204	C 165	41	55.4	596	3	CO966550	CO966550 BeE60N18F
C 93	42	56.8	675	9	CC962185	CC962185 ND1.11401	C 166	41	55.4	600	3	BI632767	BI632767 SD26634.5
C 94	42	56.8	701	9	AQ422337	AQ422337 RPCI-11-1	C 167	41	55.4	602	6	CR538947	CR538947 777044.MA
C 95	42	56.8	719	10	CW570808	CW570808 OA_ABA009	C 168	41	55.4	603	10	CW325753	CW325753 104.820.1
C 96	42	56.8	732	11	CR485728	CR485728 mtr2-154F	C 169	41	55.4	609	8	DR115048	DR115048 1432635.M
C 97	42	56.8	736	6	CB166635	CB166635 IBE603020	C 170	41	55.4	615	3	BI579387	BI579387 RE73412.5
C 98	42	56.8	748	9	CC925243	CC925243 t082m04ba	C 171	41	55.4	622	11	CR867190	CR867190 Sub scrof
C 99	42	56.8	767	9	BZ507891	BZ507891 BONQA05TR	C 172	41	55.4	623	5	BU459724	BU459724 603777648
C 100	42	56.8	769	9	BZ389085	BZ389085 EINCNC37TF	C 173	41	55.4	625	5	BQ913305	BQ913305 OHA6P07.Y
C 101	42	56.8	778	7	CR528132	CR528132 CR528132	C 174	41	55.4	629	8	BZ866700	BZ866700 CH240.224
C 102	42	56.8	778	10	CW960333	CW960333 AIAA-aab0	C 175	41	55.4	633	8	CV932315	CV932315 PWRpcm_17
C 103	42	56.8	793	2	BG822958	BG822958 602727966	C 176	41	55.4	643	9	AQ319700	AQ319700 RPCI11-10
C 104	42	56.8	803	9	CC478198	CC478198 CH240.304	C 177	41	55.4	643	10	CL409977	CL409977 RPCI44.41
C 105	42	56.8	810	7	CN056876	CN056876 Salamande	C 178	41	55.4	648	10	CG073615	CG073615 PUTMH78TD
C 106	42	56.8	817	10	CZ242866	CZ242866 AIAA-ae0	C 179	41	55.4	650	9	CC477657	CC477657 CH240.303
C 107	42	56.8	818	10	CZ301608	CZ301608 ZMMBFC077	C 180	41	55.4	651	6	CH085251	CH085251 qp82a08.9
C 108	42	56.8	819	10	CG864475	CG864475 ZMMBFC027	C 181	41	55.4	658	9	BZ328985	BZ328985 id74d03.g
C 109	42	56.8	822	10	CZ242489	CZ242489 AIAA-aag3	C 182	41	55.4	659	2	BG302136	BG302136 p803b05.y
C 110	42	56.8	823	6	CD075024	CD075024 MA3-0001U	C 183	41	55.4	659	7	CN768041	CN768041 taf66b11.
C 111	42	56.8	867	9	CC404654	CC404654 PUHUK33TB	C 184	41	55.4	659	10	CW325754	CW325754 104.820.1
C 112	42	56.8	870	7	CO880994	CO880994 BovGen_09	C 185	41	55.4	663	7	CK756213	CK756213 ltu01-2ms
C 113	42	56.8	900	10	AG839415	AG839415 Oryza sat	C 186	41	55.4	663	8	DT001550	DT001550 Mdf8000SM
C 114	42	56.8	911	9	CC100748	CC100748 CSU-K34.1	C 187	41	55.4	663	9	BH734456	BH734456 M8E5001TR
C 115	42	56.8	953	10	C2926715	C2926715 109844825	C 188	41	55.4	667	3	BI579662	BI579662 RE73729.5
C 116	42	56.8	1012	6	CA453826	CA453826 AGENCOURT	C 189	41	55.4	673	6	CF440761	CF440761 EST677106
C 117	42	56.8	1051	10	CN80051B	AL056602 Drosophil	C 190	41	55.4	674	10	CL636131	CL636131 CH243-112
C 118	42	56.8	1101	10	CN80052T	AL054596 Drosophil	C 191	41	55.4	675	2	BI171244	BI171244 RE12967.5
C 119	41	55.4	181	2	BA400648	BA400648 BH400648	C 192	41	55.4	676	9	BZ872752	BZ872752 CH240.238
C 120	41	55.4	197	9	BZ890151	BZ890151 CH240.221	C 193	41	55.4	680	9	CC712624	CC712624 OGV8046TV
C 121	41	55.4	225	9	CC491280	CC491280 CH240.324	C 194	41	55.4	684	5	BW272554	BW272554 BW272554
C 122	41	55.4	228	1	AV557572	AV557572 AV557572	C 195	41	55.4	686	11	CR832029	CR832029 GROAAA65A
C 123	41	55.4	256	2	BB202682	BB202682 BB202682	C 196	41	55.4	687	9	BZ048323	BZ048323 lk160b10.
C 124	41	55.4	275	3	BI641088	BI641088 SD24227.5	C 197	41	55.4	688	9	BH666472	BH666472 BOH254TR
C 125	41	55.4	287	7	CO975301	CO975301 BeG9N14H	C 198	41	55.4	690	11	CR499634	CR499634 mtn2-176F
C 126	41	55.4	294	1	AV160435	AV160435 AV160435	C 199	41	55.4	691	6	CB449441	CB449441 7036339.MA
C 127	41	55.4	316	1	AA429548	AA429548 zw74601.r	C 200	41	55.4	693	9	AZ856893	AZ856893 2M0161B04
C 128	41	55.4	322	9	CC543443	CC543443 CH240.425	C 201	41	55.4	696	6	CD373087	CD373087 UI-R-GOO-
C 129	41	55.4	361	2	BG576599	BG576599 BM1_55_P0	C 202	41	55.4	701	10	CL409936	CL409936 RPCI44.41
C 130	41	55.4	395	10	AB082309	AB082309 Drosophil	C 203	41	55.4	706	9	CC625602	CC625602 OGVAF59TH
C 131	41	55.4	401	9	AZ036784	AZ036784 RPCI-23-3	C 204	41	55.4	707	7	CN446983	CN446983 GUO_CDNA
C 132	41	55.4	402	9	AQ445016	AQ445016 GSSTC0656	C 205	41	55.4	713	7	CK474995	CK474995 AGENCOURT
C 133	41	55.4	410	3	BP608397	BP608397 BP608397	C 206	41	55.4	715	9	CC773066	CC773066 CH240.12H
C 134	41	55.4	412	9	AQ774176	AQ774176 HS.2050.B	C 207	41	55.4	715	10	CW588062	CW588062 OA_ABA012
C 135	41	55.4	420	9	AQ445015	AQ445015 GSSTC0646	C 208	41	55.4	717	10	CG839440	CG839440 Ynfw0054
C 136	41	55.4	448	8	W00091	W00091 TGESTzy560	C 209	41	55.4	724	10	CW914431	CW914431 RPCI42.15
C 137	41	55.4	454	6	CB077852	CB077852 h160b11.g	C 210	41	55.4	725	5	BW254522	BW254522 BW254522
C 138	41	55.4	467	6	CB1713952	CB1713952 AMGNNUC.N	C 211	41	55.4	730	5	BM399499	BM399499 BM399499
C 139	41	55.4	473	3	BI638685	BI638685 SD21039.5	C 212	41	55.4	730	9	BZ685727	BZ685727 PUBED09TD
C 140	41	55.4	477	3	BP897070	BP897070 BP897070	C 213	41	55.4	730	10	CW109686	CW109686 104.482.1
C 141	41	55.4	477	10	CL897573	CL897573 abg43b10.	C 214	41	55.4	731	7	CK478427	CK478427 AGENCOURT
C 142	41	55.4	482	10	CW891285	CW891285 RPCI42.11	C 215	41	55.4	731	9	CG522869	CG522869 CH240.371
C 143	41	55.4	496	3	BQ253042	BQ253042 eao05d06.	C 216	41	55.4	735	7	CO563972	CO563972 AGENCOURT
C 144	41	55.4	497	3	BZ782426	BZ782426 A2SP2C14	C 217	41	55.4	737	10	BX233294	BX233294 Danilo rer
C 145	41	55.4	498	2	BI236935	BI236935 RE33066.5	C 218	41	55.4	744	9	CC590916	CC590916 CH240.391
C 146	41	55.4	501	7	CN553882	CN553882 tae33g02.	C 219	41	55.4	745	9	CC585705	CC585705 CH240.382
C 147	41	55.4	513	6	CA334553	CA334553 NISC_1s10	C 220	41	55.4	754	9	BH513087	BH513087 BOGN072TR
C 148	41	55.4	519	2	BE600821	BE600821 P11_90.El	C 221	41	55.4	756	6	CF439368	CF439368 EST675713
C 149	41	55.4	525	8	DR116512	DR116512 1431631.M	C 222	41	55.4	756	6	CF444440	CF444440 EST680785
C 150	41	55.4	530	8	CV960793	CV960793 PYPCV_55	C 223	41	55.4	758	9	CC452725	CC452725 ZMMB8C034
C 151	41	55.4	532	1	AW916752	AW916752 EST348160	C 224	41	55.4	763	8	CX581776	CX581776 TTRE00255
C 152	41	55.4	534	10	CW227782	CW227782 104.667.1	C 225	41	55.4	765	5	BW492323	BW492323 BW492323
C 153	41	55.4	547	7	CN578838	CN578838 Mdfw2038F	C 226	41	55.4	771	11	CR807126	CR807126 GROAAA3BE

227	41	55.4	776	9	C0550205	CH240_434	300	40	54.1	416	3	BM625194
C 228	41	55.4	780	9	B2215419	CH230-371	301	40	54.1	424	6	CA305742
C 229	41	55.4	781	3	BI646103	603276306	302	40	54.1	425	1	AV796178
C 230	41	55.4	782	6	CF452237	EST688582	C 303	40	54.1	430	5	BY157262
C 231	41	55.4	786	3	BI685511	603308982	C 304	40	54.1	441	8	DN346359
C 232	41	55.4	792	9	BH689999	BOMWR95TF	305	40	54.1	453	10	CG919710
C 233	41	55.4	792	10	CL825062	OR_CBA004	306	40	54.1	454	9	CC470378
C 234	41	55.4	794	9	BH702782	BOHYN81TF	C 307	40	54.1	458	9	AZ153240
C 235	41	55.4	806	5	BW142868	BW142868	C 308	40	54.1	463	6	CB018032
C 236	41	55.4	813	9	CC541824	CH240_422	C 309	40	54.1	478	3	BM113082
C 237	41	55.4	818	8	DR932418	EST112395	C 310	40	54.1	483	1	AL930171
C 238	41	55.4	820	9	BZ740953	CGEDS78TC	C 311	40	54.1	486	1	AW035957
C 239	41	55.4	820	9	BZ740964	OGEDS78TC	C 312	40	54.1	496	9	BZ861735
C 240	41	55.4	822	10	BX241445	Danio rer	C 313	40	54.1	503	4	AY441057
C 241	41	55.4	826	8	CV893384	PA004F10	C 314	40	54.1	515	11	CR813199
C 242	41	55.4	826	9	CA979737	CH240_334	C 315	40	54.1	518	6	CD088748
C 243	41	55.4	830	9	AQ858416	nbe00131	C 316	40	54.1	527	6	CB401741
C 244	41	55.4	830	9	CS581393	CH240_376	C 317	40	54.1	529	2	BF648588
C 245	41	55.4	830	10	CG080883	CG080883	C 318	40	54.1	537	5	BX553026
C 246	41	55.4	844	9	BZ408632	OGAAH38TM	C 319	40	54.1	537	6	CD399561
C 247	41	55.4	850	9	BZ138972	CH230-387	C 320	40	54.1	541	2	BG387269
C 248	41	55.4	871	6	CF446583	EST682928	C 321	40	54.1	544	10	CZ290167
C 249	41	55.4	879	8	DN066339	JGI_CARD2	C 322	40	54.1	551	10	CM306882
C 250	41	55.4	880	7	CO973016	BEGON05E	C 323	40	54.1	554	1	AU159805
C 251	41	55.4	881	10	CG438303	CGVHR30TH	C 324	40	54.1	554	3	BM891913
C 252	41	55.4	882	7	CO974333	BeG9ON04C	C 325	40	54.1	555	9	CC647255
C 253	41	55.4	884	10	CG052650	PUPRO76TD	C 326	40	54.1	555	9	CC647261
C 254	41	55.4	885	10	CZ228814	ZMBBF0035	C 327	40	54.1	562	9	AZ030534
C 255	41	55.4	892	2	BG830768	602767467	C 328	40	54.1	570	5	BU720959
C 256	41	55.4	892	9	CS973654	ZUAHC33TH	C 329	40	54.1	573	10	CM106027
C 257	41	55.4	902	7	CO975640	CO975640	C 330	40	54.1	577	10	CM106027
C 258	41	55.4	918	11	CNS06X17	AL419153 T3 end of	C 331	40	54.1	577	2	BF647021
C 259	41	55.4	937	10	CG230500	OGSCK34TC	C 332	40	54.1	577	6	CA196584
C 260	41	55.4	957	10	CL090564	ISB1-17N2	C 333	40	54.1	582	7	CO851815
C 261	41	55.4	969	10	CL047142	CH216-65M	C 334	40	54.1	586	5	BX782673
C 262	41	55.4	978	10	CG417187	ZMBB0002	C 335	40	54.1	588	11	P957L
C 263	41	55.4	978	10	CL480496	SAIL_325	C 336	40	54.1	590	1	AJ758771
C 264	41	55.4	1005	10	CL235468	CL235468	C 337	40	54.1	590	7	CO834889
C 265	41	55.4	1025	9	CC240632	CH261-120	C 338	40	54.1	591	6	CD339355
C 266	41	55.4	1061	10	CL456998	CL456998	C 339	40	54.1	595	6	CB174129
C 267	41	55.4	1177	9	CC242258	CH261-123	C 340	40	54.1	597	7	CO908853
C 268	41	55.4	1239	5	BU501641	AGENCOURT	C 341	40	54.1	606	2	BF647562
C 269	41	55.4	1242	10	DU021821	DU021821 1524 Toma	C 342	40	54.1	609	2	BF649961
C 270	41	55.4	1318	3	BQ233334	BQ233334	C 343	40	54.1	611	11	DE025279
C 271	40	54.1	166	2	BB433377	BB433377	C 344	40	54.1	613	10	CM641416
C 272	40	54.1	222	1	BB163045	BB163045	C 345	40	54.1	619	10	CM042643
C 273	40	54.1	240	5	CG3262	CG3262 Yuji	C 346	40	54.1	625	1	AU301922
C 274	40	54.1	241	2	BB299854	BB299854	C 347	40	54.1	629	11	CR484779
C 275	40	54.1	246	1	AW264700	AW264700 xq34a10.x	C 348	40	54.1	630	2	BG587363
C 276	40	54.1	271	10	CM887251	CM887251	C 349	40	54.1	630	10	AG157590
C 277	40	54.1	272	3	BP581431	BP581431	C 350	40	54.1	633	2	BG448424
C 278	40	54.1	282	7	CV285296	tak43e08	C 351	40	54.1	635	3	BI928201
C 279	40	54.1	287	2	BB536900	BB536900	C 352	40	54.1	638	9	BZ222971
C 280	40	54.1	291	10	AL766882	AL766882	C 353	40	54.1	641	11	DE055488
C 281	40	54.1	294	1	BB125246	BB125246	C 354	40	54.1	651	2	BF648419
C 282	40	54.1	295	2	BB306396	BB306396	C 355	40	54.1	651	7	CM791572
C 283	40	54.1	295	11	DE108365	DE108365	C 356	40	54.1	651	7	CV007828
C 284	40	54.1	301	8	CV937853	PMDCm0_00	C 357	40	54.1	651	11	DE065354
C 285	40	54.1	308	2	BG210196	BG210196	C 358	40	54.1	653	3	BU733355
C 286	40	54.1	309	7	CV285088	CV285088	C 359	40	54.1	654	10	CM434174
C 287	40	54.1	315	10	CG881506	CG881506	C 360	40	54.1	654	10	CM496043
C 288	40	54.1	323	2	BB459727	BB459727	C 361	40	54.1	655	2	BG356833
C 289	40	54.1	335	1	AV790772	AV790772	C 362	40	54.1	657	11	CR341463
C 290	40	54.1	342	2	BG187405	BG187405	C 363	40	54.1	659	6	CD074045
C 291	40	54.1	359	1	AI152043	ud19a09.r	C 364	40	54.1	661	6	CD808984
C 292	40	54.1	360	5	CG2787	CG2787	C 365	40	54.1	663	2	BG594900
C 293	40	54.1	360	8	D35007	C65473 Yuji	C 366	40	54.1	664	9	BZ887487
C 294	40	54.1	360	5	CG5473	CG5473	C 367	40	54.1	665	8	CM351977
C 295	40	54.1	372	5	CG3047	CG3047	C 368	40	54.1	666	9	CC570383
C 296	40	54.1	372	9	BH732613	BH732613	C 369	40	54.1	668	2	BF646345
C 297	40	54.1	391	8	DR108795	DR108795	C 370	40	54.1	668	10	CM190803
C 298	40	54.1	407	9	AQ882677	HS_5431_B	C 371	40	54.1	669	3	BU491885
C 299	40	54.1	409	9	CC887019	SALK_1493	C 372	40	54.1	671	9	BM746043

BM625194 170006874
CA305742 EST000012
AV796178 AV796178
BY157262 BY157262
DN346359 LIB3551-0
CG919710 MBRFC38TR
CC470378 CH240_143
AZ153240 SP_0009_B
CB018032 pgnic.pk0
BM113082 EST50618
AL930171 AL930171
AW035957 EST282816
BZ861735 CH240_246
AY441057 Armigeres
CR813199 GROAAA37C
CD088748 MCL-0053T
CB401741 OSTF198F2
BF648588 NF049B08E
BX553026 BX53026
CD399561 Gm_CK2124
BG387269 602456016
CZ290167 CP68d01.r
CM306882 104_792_1
AU159805 AU159805
BM891913 sam4ea08
CC647255 OGUDU43TH
CC647261 OGUDU43TV
AZ030534 RPCI-23-2
A2030534 RPCI-23-2
BU720959 S7M2BIB2
BF649410 NF079F01E
CM106027 104_475_1
BF647021 NF092D12E
CA196584 SCBFAD109
CO851815 LM_SH5_00
BX782673 BX782673
AU390660 Letahmani
AJ758771 AJ758771
CO834889 LM_GH5_00
CD339355 STPu536
CB174129 OR_2034H0
CO908853 BJ2037A0
BF647562 NF011E08E
BF649961 NF086A08E
DE025279 Brachios
CM641416 OA_ABA017
CM042643 104_276_1
AU301922 AU301922
CR484779 mcn2-152B
BG587363 EST489135
AG157590 Pan_tregl
BG448424 NF024E05E
BI928201 EST548090
BZ222971 ia72d08.g
DE055488 Oryzias_1
BF648419 NF047B12E
CM791572 4126318_B
CV007828 zfi01-11m
DE065354 Oryzias_1
BU733355 BJ733355
CM434174 f8bb001f1
CM496043 f8bb001f2
BG356833 OV2_10_E1
CR341463 mte1-73F1
CD074045 MA3-0001U
CD808984 LCA03D12C
BG594900 EST493578
BZ887487 CH240_255
CM351977 s13GNF62A
CC570383 CH240_445
BF646345 NF068H08E
CM190803 104_612_1
BU491885 BU491885
BM746043 qn15d10.b

373	40	54.1	674	7	CO006720	CO006720 EST795055	C 446	40	54.1	863	10	CZ282271	CZ282271 cp22b01.f
374	40	54.1	675	8	DN544035	DN544035 1392928 M	447	40	54.1	867	10	CG216672	CG216672 CG0BT09TH
375	40	54.1	676	8	DN544035	DN544035 1392928 M	448	40	54.1	870	6	CF694500	CF694500 CCAPO52TR
376	40	54.1	677	10	DN544035	DN544035 1392928 M	449	40	54.1	871	9	BZ286510	BZ286510 PUGB507DR
377	40	54.1	678	10	AX1057627	AX1057627 Danlo rer	450	40	54.1	878	11	CR797350	CR797350 GR0AAA15D
378	40	54.1	679	1	AJ779001	AJ779001 AJ779001	C 451	40	54.1	880	6	CB565951	CB565951 AGENCOURT
379	40	54.1	680	8	AX1057627	AX1057627 Danlo rer	452	40	54.1	883	9	BZ233810	BZ233810 CH230-427
380	40	54.1	681	8	AX1057627	AX1057627 Danlo rer	453	40	54.1	883	10	CG298777	CG298777 CG0BM77TH
381	40	54.1	682	10	CO063306	CO063306 104_309_1	C 454	40	54.1	886	9	CG551098	CG551098 CH240_436
382	40	54.1	683	10	CO063306	CO063306 104_309_1	C 455	40	54.1	886	9	CG551098	CG551098 CH240_436
383	40	54.1	684	5	BU223390	BU223390 603947711	C 456	40	54.1	890	7	CO930738	CO930738 AGENCOURT
384	40	54.1	685	6	CF685373	CF685373 CCAAW59TR	C 457	40	54.1	892	10	CL075513	CL075513 CH216-136
385	40	54.1	686	7	CO868533	CO868533 Mdbb5012a	C 458	40	54.1	891	6	DR942313	DR942313 EST113385
386	40	54.1	687	1	AV165081	AV165081 AV165081	C 459	40	54.1	901	6	CF686776	CF686776 CCAGP19TF
387	40	54.1	688	7	CO010754	CO010754 EST799089	C 460	40	54.1	924	7	CK294694	CK294694 EST757408
388	40	54.1	689	10	AG186214	AG186214 Pan trogl	C 461	40	54.1	925	9	BZ26507	BZ26507 PUGB507DR
389	40	54.1	690	7	CO006721	CO006721 EST795056	C 462	40	54.1	926	6	CD247523	CD247523 AGENCOURT
390	40	54.1	691	3	BJ495021	BJ495021 BJ495021	C 463	40	54.1	926	11	CNS04T3M	CNS04T3M Tetraodon
391	40	54.1	692	10	CG247065	CG247065 1023768 R	C 464	40	54.1	928	10	CZ963113	CZ963113 161109 To
392	40	54.1	693	10	CG887487	CG887487 ZMBBb035	C 465	40	54.1	929	6	CF694494	CF694494 CCAPO52TR
393	40	54.1	694	10	AX171555	AX171555 Danlo rer	C 466	40	54.1	937	7	CK020827	CK020827 AGENCOURT
394	40	54.1	695	9	AZ524701	AZ524701 235PDD04	C 467	40	54.1	946	7	CK285936	CK285936 EST748658
395	40	54.1	696	5	BY711650	BY711650 BY711650	C 468	40	54.1	947	10	DU007743	DU007743 30859 To
396	40	54.1	697	8	BH540661	BH540661 BQWH69TF	C 469	40	54.1	947	10	DU008473	DU008473 30859 To
397	40	54.1	698	7	CV898705	CV898705 PB014C5 m	C 470	40	54.1	953	10	CZ320499	CZ320499 ZMBBb0022
398	40	54.1	699	8	CF678843	CF678843 CCAFP40TR	C 471	40	54.1	966	10	DU024894	DU024894 5164 Toma
399	40	54.1	700	10	AX165720	AX165720 Danlo rer	C 472	40	54.1	966	10	CG328691	CG328691 OGXCN71TV
400	40	54.1	701	8	BH730576	BH730576 BOMB711TR	C 473	40	54.1	974	9	CZ333849	CZ333849 CH261-66D
401	40	54.1	702	8	DR912427	DR912427 EST110396	C 474	40	54.1	978	6	CF695020	CF695020 CCAPO52TR
402	40	54.1	703	9	BZ672171	BZ672171 PUBB711TD	C 475	40	54.1	979	5	BU507239	BU507239 AGENCOURT
403	40	54.1	704	10	CG905329	CG905329 RPI142_14	C 476	40	54.1	982	2	BG024012	BG024012 602303226
404	40	54.1	705	8	DN096851	DN096851 JGI CABE7	C 477	40	54.1	1008	11	CNS05DK6	CNS05DK6 Tetraodon
405	40	54.1	706	7	CO010755	CO010755 EST799090	C 478	40	54.1	1009	9	CC240484	CC240484 CH261-132
406	40	54.1	707	9	BZ316462	BZ316462 1872808.D	C 479	40	54.1	1029	10	CL023150	CL023150 CH216-281
407	40	54.1	708	3	CF695024	CF695024 CCAPO52TR	C 480	40	54.1	1043	10	CL037439	CL037439 CH216-43E
408	40	54.1	709	6	BI824316	BI824316 603040733	C 481	40	54.1	1046	10	CM001528	CM001528 ZMBBb0000
409	40	54.1	710	10	CL864594	CL864594 TM1-GSS00	C 482	40	54.1	1101	10	CNS00DV5	CNS00DV5 Drosophila
410	40	54.1	711	6	CD079787	CD079787 MA3-99990	C 483	40	54.1	1116	9	CC192052	CC192052 CH261-6M2
411	40	54.1	712	6	CF704703	CF704703 CCAPO52TR	C 484	40	54.1	1126	8	DR133510	DR133510 47512731
412	40	54.1	713	10	AG463491	AG463491 Nus muscu	C 485	40	54.1	1126	8	DR133510	DR133510 47512731
413	40	54.1	714	10	AL119229	AL119229 DKFP2761C	C 486	40	54.1	1199	9	CC194132	CC194132 CH261-560
414	40	54.1	715	8	DR941233	DR941233 EST113277	C 487	40	54.1	1262	3	BI689878	BI689878 603310710
415	40	54.1	716	7	BH341457	BH341457 CH230-89L	C 488	40	54.1	1346	10	CL646332	CL646332 CH213-113
416	40	54.1	717	9	CC494353	CC494353 CH240_329	C 489	40	54.1	1390	8	DN716204	DN716204 CNB111-80
417	40	54.1	718	6	CF680634	CF680634 CCAPO52TR	C 490	40	54.1	1508	3	BM803190	BM803190 AGENCOURT
418	40	54.1	719	1	AJ793883	AJ793883 AJ793883	C 491	40	54.1	1666	8	DN987266	DN987266 MSU_28F_2
419	40	54.1	720	11	CR823369	CR823369 GROAAA53C	C 492	40	54.1	210	1	BB000247	BB000247 BB000247
420	40	54.1	721	10	CM704704	CM704704 AIAA-aab9	C 493	40	54.1	467	9	AQ680174	AQ680174 HS_5470_A
421	40	54.1	722	2	BG585003	BG585003 EST486765	C 494	40	54.1	566	9	AZ499990	AZ499990 IM0338B07
422	40	54.1	723	10	CM523052	CM523052 OP_Ba001	C 495	40	54.1	596	5	BU781385	BU781385 kb18f05.Y
423	40	54.1	724	9	CG969205	CG969205 MBEDC56TR	C 496	40	54.1	804	8	DR779203	DR779203 BAAC-PNPI
424	40	54.1	725	9	BZ394032	BZ394032 EINBU57TR	C 497	40	54.1	804	8	DR779203	DR779203 BAAC-PNPI
425	40	54.1	726	6	CF695688	CF695688 CCAHCl17TR	C 498	40	54.1	121	9	CC040298	CC040298 3591_I_13
426	40	54.1	727	5	BQ505997	BQ505997 EST613412	C 499	40	54.1	156	1	AA168211	AA168211 ms51f11.r
427	40	54.1	728	6	CF680650	CF680650 CCAPO52TR	C 500	40	54.1	159	1	AA168211	AA168211 ms51f11.r
428	40	54.1	729	8	CF508822	CF508822 JGI_XZG65							AA371700 EST83571
429	40	54.1	730	6	CF676373	CF676373 CCAPO52TR							
430	40	54.1	731	10	CZ281891	CZ281891 cp19e03.r							
431	40	54.1	732	9	BH725214	BH725214 BOMBA23TR							
432	40	54.1	733	8	DR500671	DR500671 WS02921.C							
433	40	54.1	734	6	CB635261	CB635261 OS1EB15H							
434	40	54.1	735	6	CB943164	CB943164 AGENCOURT							
435	40	54.1	736	10	DU069910	DU069910 137333 To							
436	40	54.1	737	6	CF695682	CF695682 CCAHCl17TF							
437	40	54.1	738	6	CF686782	CF686782 CCAGP19TR							
438	40	54.1	739	8	DR915632	DR915632 EST110717							
439	40	54.1	740	11	CR824747	CR824747 GROAAA55C							
440	40	54.1	741	8	CD049473	CD049473 SV6_39F08							
441	40	54.1	742	8	DN981657	DN981657 SV6_39F08							
442	40	54.1	743	10	CG238888	CG238888 OG0BN77TH							
443	40	54.1	744	10	AG886962	AG886962 Oryza sat							
444	40	54.1	745	8	CK391333	CK391333 JGI_XZT38							
445	40	54.1	746	8	DN981658	DN981658 SV6_39F09							

ALIGNMENTS

RESULT 1
LOCUS
DEFINITION
CNSOABAJ/c
CNSOABAJ
541 bp mRNA linear HTC 06-FEB-2004
Arabidopsis thaliana Full-length cDNA Complete sequence from clone
GSLTSIL212F2 of Silique of strain col-0 of Arabidopsis thaliana
(thale cress)
BX817426
GI:42471008
BX817426
HTC; GSIT cDNA
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

REFERENCE 1 (bases 1 to 541)
 AUTHORS Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C., Menard,M., Cruaud,C., Quetier,P., Scarpelli,C., Schachter,V., Temple,G., Caboche,M., Weissenbach,J. and Salanoubat M.
 TITLE Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
 A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 541)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
 COMMENT The sequences are based on single pass reads.
 Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.
 Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.
 URGV INRA : Clepet C., Caboche M.
 Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap.
 http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full length
 http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.
 FEATURES
 source
 Location/Qualifiers
 1..541
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /clone="GSLTSL1212F12"
 /tissue_type="Silique"
 /ecotype="Col-0"
 /plasmid="pCMVSPORT_6"

ORIGIN
 Alignment Scores:
 Pred. No.: 65.4 Length: 541
 Score: 47.00 Matches: 9
 Percent Similarity: 90.00% Conservative: 0
 Best Local Similarity: 90.00% Mismatches: 1
 Query Match: 63.51% Indels: 0
 DB: 4 Gaps: 0

US-10-787-382-20 (1-15) x CNS0ABAJ (1-541)
 Qy 5 AenProMetAsnArgLeuValAlaGluThr 14
 |||||
 Db 489 AATCCCATGACAGATTGGCAGCGAGACC 460
 |||||

RESULT 2
 AI002749/c AI002749 376 bp mRNA linear EST 09-JUN-1998
 LOCUS am98a04.s1 Stratagene schizo brain S11 Homo sapiens cDNA clone
 DEFINITION IMAGE:1684110 3', similar to TR:Q13500 Q13500 DYSTROBREVIN-EPSILON.
 ; mRNA sequence.
 ACCESSION AI002749
 VERSION AI002749.1 GI:3203163
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 376)
 AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S., Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
 WASHU-NCI human EST project
 Unpublished (1997)

COMMENT Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Trace considered overall poor quality
 Possible reversed clone: similarity on wrong strand
 Seq primer: -40ml3 fwd. ET from Amersham
 High quality sequence stop: 1.
 Location/Qualifiers
 1..376
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:1684110"
 /sex="male"
 /tissue_type="schizophrenic brain S-11 frontal lobe"
 /dev_stage="34 years old"
 /lab_hosts="SOLR (kanamycin resistant)"
 /clone_lib="Stratagene schizo brain S11"
 /note="Vector: Bluescript SK-; Site_1: EcoRI; Library constructed from S-11 frontal lobe, male, 34 years old, 50% caucasian, 50% Aleutian. Schizophrenic suicide. Random primed into EcoRI site of ZAP II Vector. Mass excised. Custom library. Avg insert length 1.4kb. Material obtained by Johnston N., Torrey, E.F., Yolken R., and the Stanley Neuropathology Consortium - Analysis of RNAs from the Brains of Individuals with Psychiatric Diseases (Unpublished) Stanley Neurovirology Laboratory, Johns Hopkins School of Medicine, Baltimore MD."

ORIGIN
 Alignment Scores:
 Pred. No.: 68.6 Length: 376
 Score: 46.00 Matches: 9
 Percent Similarity: 76.92% Conservative: 1
 Best Local Similarity: 69.23% Mismatches: 3
 Query Match: 62.16% Indels: 0
 DB: 1 Gaps: 0

US-10-787-382-20 (1-15) x AI002749 (1-376)
 Qy 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGlu 13
 |||||
 Db 272 TTTCGCTAGAAACCCCTTTTCAGAGGTGGTGGAGGAA 234
 |||||

RESULT 3
 CO192024
 LOCUS CO192024 565 bp mRNA linear EST 21-JUN-2004
 DEFINITION EC32618.5prime Exelixis FlyTag ML01 pSport-Tag21 Drosophila melanogaster cDNA clone EC32618 5, mRNA sequence.
 ACCESSION CO192024
 VERSION CO192024.1 GI:49003199
 KEYWORDS EST.
 SOURCE Drosophila melanogaster (fruit fly)
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 REFERENCE 1 (bases 1 to 565)
 AUTHORS Chen,F., Lagueux,M., Cheung,L.M., Chong,A., Goldschmidt,S., Hubsain,S., Laufer,A., Oliva,J., Park,C., Wong,M., Amundsen,C., Orton,A., Shao,A., Platt,D. and Swimmer,C.
 TITLE Exelixis FlyTag EST Project ML01 Library
 JOURNAL Unpublished (2004)
 COMMENT Contact: Stapleton, M.
 BDGP
 Lawrence Berkeley National Lab
 One Cyclotron Rd, Berkeley, CA 94720, USA

```

Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
Plate: EC.326 row: B column: 6
High quality sequence stop: 504.
Location/Qualifiers
1. .565
/morganism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="EC32618"
/dev_stage="3rd instar larva"
/clone_lib="Exelixis FlyTag ML01 pSport1-Tag21"
Note="Organ: fat body; Vector: pSport1-Tag21; Site 1:
Not1; Site 2: XhoI; RNA was isolated from fat body from
3rd instar larva challenged with gram-/ - bacteria. cDNA
was oligodT primed."

ORIGIN
Alignment Scores:
Pred. No.: 109 Length: 565
Score: 46.00 Matches: 8
Percent Similarity: 83.3% Conservative: 2
Best Local Similarity: 66.6% Mismatches: 2
Query Match: 62.1% Indels: 0
DB: 7 Gaps: 0

US-10-787-382-20 (1-15) x C0192024 (1-565)

QY 1 PheAlaValGluAsnProMetAsnArgLeuValAla 12
Db 177 TTTTCTGCTGCAACCCATTAAACCGATTGATCGCG 212

RESULT 4
CR821531
LOCUS
DEFINITION
CR821531
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
CR821531
GROAAASODE10FM1 INRA BAC Bos taurus genomic clone INRA8_810F01, DNA
sequence, genomic survey sequence.
CR821531.1 GI:52757619
GSS.
Bos taurus (cow)
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Bovinae; Bos.
1 (bases 1 to 874)
Eggen, A., Schibler, L. and Roy, A.
Bovine BAC End Sequences from the INRA bovine BAC library
Unpublished
2 (bases 1 to 874)
Genoscope.
Direct Submission
Submitted (20-SEP-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
Contact: Andre Eggen
Department of Animal Genetics - IGBIC
INRA
78350 Jouy-en-Josas, France
Tel: 33 1 34 65 24 24
Fax: 33 1 34 65 24 78
Email: eggen@jouy.inra.fr
Clones are derived from the INRA bovine BAC library
(http://locus.jouy.inra.fr/fpc/cattle_bac_map.htm). For BAC library
availability, please contact Andre Eggen (eggen@jouy.inra.fr). This
work was undertaken as part of the International Bovine BAC
Mapping Consortium (IBBMC) by INRA (Jouy-en-Josas) and Genoscope
(Evry) Plate: 810 row: F column: 01
Seq primer: M13 Forward
Class: BAC ends.
Location/Qualifiers
1. .874
/organism="Bos taurus"

```

```

/mol_type="genomic DNA"
/strain="breed: Holstein"
/db_xref="taxon:9913"
/clone="INRA8_810F01"
/sex="Male"
/cell_type="fibroblast"
/clone_lib="INRA bovine BAC"
Note="Vector: pBelobAC11; Site 1: HindIII; Holstein bull;
INRA Bovine BAC library (Male) produced by Andre Eggen
Genoscope sequence ID : GROAAASODE10FM1"

ORIGIN
Alignment Scores:
Pred. No.: 180 Length: 874
Score: 46.00 Matches: 8
Percent Similarity: 78.5% Conservative: 3
Best Local Similarity: 57.1% Mismatches: 3
Query Match: 62.1% Indels: 0
DB: 11 Gaps: 0

US-10-787-382-20 (1-15) x CR821531 (1-874)

QY 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThr 14
Db 722 TTTTCTGCTGGAATCCTATGAACAGAGAATCTGGCAGACT 763

RESULT 5
AV222732/c
LOCUS
DEFINITION
AV222732
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
AV222732
RIKEN full-length enriched, 18 days pregnant, placenta and
extra embryonic tissue Mus musculus cDNA clone 3830403007 3', mRNA
sequence.
AV222732
AV222732.1 GI:6171909
EST.
Mus musculus
Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
1 (bases 1 to 256)
Konno, H., Aizawa, K., Akahira, S., Akiyama, J., Carninci, P., Endo, T.,
Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Hori, F.,
Ishii, Y., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I.,
Kai, C., Kawai, J., Kikuchi, N., Kojima, Y., Koya, S., Kusakabe, M.,
Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y.,
Owa, C., Ozawa, Y., Saito, H., Sano, M., Sato, K., Shibata, K.,
Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara, Y.,
Suzuki, H., Suzuki, H., Takahashi, F., Tatenoe, M., Tomine, N.,
Tsunoda, Y., Watahiki, A., Watanabe, S., Yamamura, T., Yasunishi, A.,
Yokota, T., Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
RIKEN Mouse ESTs (Konno, H., et al. 1999)
Unpublished (1999)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gs.c.riken.jp, URL: http://genome.gsc.riken.jp/
Sasaki, N., Izawa, M., Watahiki, M., Ozawa, K., Tanaka, T., Yoneda, Y.,
Matsura, S., Carninci, P., Muramatsu, M., Okazaki, Y. and
Hayashizaki, Y.
Transcriptional sequencing: A method for DNA sequencing using RNA
polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)
Itoh, M., Kitanai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,
Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki,
Y. and Hayashizaki, Y.
Automated filtration-based high-throughput plasmid preparation
system. Genome Res. 9 (5), 463-470 (1999)
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303,

```


FEATURES

```

/tissue_type="placenta and extra embryonic tissue"
/dev_stage="18 days pregnant, adult"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 18 days pregnant, placenta and extra embryonic tissue"
/note="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGATCATGAGCTCTTTTCTTTTCTT 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 50.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATCTTCAGTTAAATTAATTCCTCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from LambdaFLC 1"

```

Yy 19 TTTACCATTTCTAAATCCTCTGAAATCGATTTGTTCGGCTCCACTGTA 63

US-10-787-382-20 (1-15) x AV222732 (1-256)

RESULT 6

ACCESSION DN847029

KEYWORDS EST. BOG FARMING (cont.)

CREATISM **DOS CAULUS**
Eukaryota: Metazoa

Pecora; Bovidae;

AUTHORS	TITLE
Olsaker, I. and L.	Production of ES

JOURNAL
Unpublished (200

Dept. of Basic S

P.O. BOX 8146 DE
TOL: 147 3386477

2025/11/22 11:53

FEATURES	Location/Qualifiers
source	1. : 438

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

Alignment Scores:

US-10-787-382-20 (1-15) x DN847029 (1-438)

19 TTTTACCATTTCTAATACCTCTGAATCGAATTTGTCCGGCTCCACTGTA 63

CN755695/c

DEFINITION

ACCESSION
VERSION

SOURCE

RESEARCH

CONCLUSION

COMMENT

020117822

0
4
5
0
7

```

/clone="ID0AA16CD02"
/tissue_type="whole insect"
/dev_stage="nymphs and adults (parthenogenetic females)"
/lab_host="XLI-Blue"
/clone_lib="ApMS"
/notes="Vector: pBS-SK minus; Site 1: EcoRI; Site 2: XhoI;
Sample name: ID0AAA; Plant growth place: Department of
Ecology & Evolutionary Biology, Princeton University;
Soil conditions: Soil; Sowing date: 01/06/1999;
Harvesting date: 01/06/1999; Stress date: no stress;
Description: Aphids inoculated on one-week old Vicia faba
under non-sterile conditions. All parthenogenetic stages
and both winged and wingless adults were collected for
library construction.; experimental condition: long
photoperiod (16-hr light/8-hr dark at 18 c)"

```

ORIGIN

Alignment Scores:

```

Pred. No.: 148 Length: 490
Score: 45.00 Matches: 6
Percent Similarity: 84.62% Conservatives: 5
Best Local Similarity: 46.15% Mismatches: 2
Query Match: 60.81% Indels: 0
DB: 7 Gaps: 0

```

US-10-787-382-20 (1-15) x CN755695 (1-490)

```

QY 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGlu 13
||| :|||:|||||:|||||:|||||:|||||:|||||:
Db 45 TTTATCATCGATAACCCCACTCAATAGATTATCATAGAA 7

```

RESULT 8

CF587487/c

LOCUS

DEFINITION USDA-FP 120800-038 Acyrthosiphon pisum, Pea Aphid Acyrthosiphon pisum cDNA clone WHAP- 005_E05 5', mRNA sequence.

ACCESSION

CF587487

VERSION

CF587487.1

KEYWORDS

EST.

SOURCE

ACyrthosiphon pisum (pea aphid)

ORGANISM

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidiformes;

Aphidoidea; Aphididae; Macrosiphini; Acyrthosiphon.

1 (bases 1 to 504)

Hunter,W.B., Dang,P.M., Stern,D. and Lapointe,S.L.

Gene expression in Acyrthosiphon pisum (Hemiptera: Aphididae)

Unpublished (2004)

Contact: Wayne Hunter; Phat Dang

US. Horticultural Research Lab, Ft. Pierce, FL

USDA-ARS

2001 South Rock Rd., Ft. Pierce, FL 34945, USA

Tel: (772) 462-5898, (772) 462-5940

Fax: (772) 462-5986

Email: Whunter@usda.ars.usda.gov

Seq primer: T3 Primer.

Location/Qualifiers

1..504

/organism="Acyrthosiphon pisum"

/mol_type="mRNA"

/db_xref="taxon:7029"

/clone="WHAP-005_E05"

/sex="Mixed population"

/tissue_type="whole body, nymphs and adults"

/dev_stage="All"

/lab_host="XLI-Blue"

/clone_lib="Acyrthosiphon pisum, Pea Aphid"

/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:

XhoI; Acyrthosiphon pisum colonies. Library by Srin

Kambampati; Peter Dearden; David Stern, Department of

Ecology and Evolutionary Biology, Princeton University,

NJ. (609) 258-0759. A high quality EST with at least 100

contiguous bases at Trace Tuner score of 20 or better.

ORIGIN

Alignment Scores:

```

Pred. No.: 152 Length: 504
Score: 45.00 Matches: 6
Percent Similarity: 84.62% Conservatives: 5
Best Local Similarity: 46.15% Mismatches: 2
Query Match: 60.81% Indels: 0
DB: 6 Gaps: 0

```

US-10-787-382-20 (1-15) x CF587487 (1-504)

```

QY 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGlu 13
||| :|||:|||||:|||||:|||||:|||||:|||||:
Db 174 TTTATCATCGATAACCCCACTCAATAGATTATCATAGAA 136

```

RESULT 9

CN753056/c

LOCUS

DEFINITION CN753056 539 bp mRNA linear EST 19-MAY-2004 AphL3LD-XI-F6 AphL3LD Acyrthosiphon pisum cDNA clone AphL3LDXIF6 5', mRNA sequence.

ACCESSION

CN753056

VERSION

CN753056.1

KEYWORDS

EST.

SOURCE

Acyrthosiphon pisum (pea aphid)

ORGANISM

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidiformes;

Aphidoidea; Aphididae; Macrosiphini; Acyrthosiphon.

1 (bases 1 to 539)

Hunter,W., Martinez-Torres,D., Rahbe,Y., Sabater-Munoz,B.,

Stern,D., Tagu,D. and Wincker,P.

An expressed sequence tags database for the pea aphid Acyrthosiphon

pisum

Unpublished (2004)

Contact: D. Tagu

INRA Rennes

UMR BiO3P, BP 35327, F-35653 Le Rheu Cedex France

Tel: +33.2.23.48.51.65

Fax: +33.2.23.48.51.50

Risk of contamination by bacterial sequences from obligatory

(Buchnera) or facultative endosymbionts.

PCR PRIMERS

FORWARD: GCCGATTAACCTTCGTATAGCA

Plate: XI row: F column: 6.

Location/Qualifiers

1..539

/organism="Acyrthosiphon pisum"

/mol_type="mRNA"

/cultivar="yr2"

/db_xref="taxon:7029"

/clone="AphL3LDXIF6"

/tissue_type="head"

/dev_stage="third instar nymph (L3)"

/lab_host="TOPI0"

/clone_lib="AphL3LD"

/note="Vector: pDNR-LIB; Site 1: SfiI; Site 2: SfiI;

Sample name: AphL3LD; Plant growth place: INRA-Rennes,

UMR BiO3P, BP 35327, 35653 Le Rheu cedex, France; Soil

conditions: peat; Sowing date: 18/01/2003; Harvesting

date: 03/02/2003; Stress date: no stress; Description:

aphids inoculated on one-week old Vicia faba germinations

under non sterile conditions.; experimental condition:

long photoperiod (16-hr light/8-hr dark at 18 c)"

Sequencing of clones by Dr. PM Dang, US. Horticultural Research Lab, Ft. Pierce, FL. (772) 462- 5940."

ORIGIN

Alignment Scores:

```

Pred. No.: 165 Length: 539
Score: 45.00 Matches: 6
Percent Similarity: 84.62% Conservatives: 5
Best Local Similarity: 46.15% Mismatches: 2

```

Query Match:	60.81%	Indels:	0
DB:	7	Gaps:	0
US-10-787-382-20 (1-15) x CN753056 (1-539)			
QY	1	PheAlaValGluAsnProMetAsnArgLeuValAlaGlu	13
DB	377	TTTATCATCGAATACCCCACTCAATAGAAATTATCATAGAA	339
RESULT 10			
CN754327/c			
LOCUS	CN754327	625 bp	linear
DEFINITION	ID0AAA12CG12RM1 ApMs Acyrthosiphon pisum cDNA clone ID0AAA12CG12		EST 19-MAY-2004
ACCESSION	CN754327		
VERSION	CN754327.1	GI:47519324	
KEYWORDS	EST.		
SOURCE	Acyrthosiphon pisum (pea aphid)		
ORGANISM	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidiformes; Aphidoidea; Aphididae; Macrosiphini; Acyrthosiphon.		
REFERENCE	1 (bases 1 to 625)		
AUTHORS	Hunter,W., Martinez-Torres,D., Rahbe,Y., Sabater-Munoz,B., Stern,D., Tagu,D. and Wincker,P.		
TITLE	An expressed sequence tags database for the pea aphid Acyrthosiphon pisum		
JOURNAL	Unpublished (2004)		
COMMENT	Contact: D. Tagu		
INRA Rennes			
UMR BIO3P	BP 35327, F-35653 Le Rheu Cedex France		
Tel:	+33.2.23.48.51.65		
Fax:	+33.2.23.48.51.50		
Risk of contamination by bacterial sequences from obligatory (Buchnera) or facultative endosymbionts. These sequences were obtained in the frame of the International Consortium of Aphid Genomics in collaboration with Genoscope			
PCR Primers	FORWARD: CAGGAACAGCTATGACC		
Plate:	12 row: G column: 12.		
FEATURES			
source	1. 625		
	Location/Qualifiers		
	/organism="Acyrthosiphon pisum"		
	/mol_type="mRNA"		
	/cultivar="developmentstage"		
	/db_xref="taxon:7029"		
	/clone="ID0AAA12CG12"		
	/issue_type="whole insect"		
	/dev_stage="nymphs and adults (parthenogenetic females)"		
	/lab_host="XLI-Blue"		
	/clone_lib="ApMS"		
	/note="Vector: PBS-SK minus; Site 1: EcoRI; Site 2: XhoI; Sample name: ID0AAA ; Plant growth place: Department of Ecology & Evolutionary Biology, Princeton University ; Soil conditions: Soil ; Sowing date: 01/06/1999 ; Harvesting date: 01/06/1999 ; Stress date: no stress ; Description: Aphids inoculated on one-week old Vicia faba under non-sterile conditions. All parthenogenetic stages and both winged and wingless adults were collected for library construction. ; experimental condition: long photoperiod (16-hr light/8-hr dark at 18 c)"		
ORIGIN			
Alignment Scores:			
Pred. No.:	195	Length:	625
Score:	45.00	Matches:	6
Percent Similarity:	84.62%	Conservative:	5
Best Local Similarity:	46.15%	Mismatches:	2
Query Match:	60.81%	Indels:	0
DB:	7	Gaps:	0
US-10-787-382-20 (1-15) x CN754327 (1-625)			
QY	1	PheAlaValGluAsnProMetAsnArgLeuValAlaGlu	13
DB	377	TTTATCATCGAATACCCCACTCAATAGAAATTATCATAGAA	339
RESULT 10			
CN754327/c			
LOCUS	CN754327	625 bp	linear
DEFINITION	ID0AAA12CG12RM1 ApMs Acyrthosiphon pisum cDNA clone ID0AAA12CG12		EST 19-MAY-2004
ACCESSION	CN754327		
VERSION	CN754327.1	GI:47519324	
KEYWORDS	EST.		
SOURCE	Acyrthosiphon pisum (pea aphid)		
ORGANISM	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidiformes; Aphidoidea; Aphididae; Macrosiphini; Acyrthosiphon.		
REFERENCE	1 (bases 1 to 625)		
AUTHORS	Hunter,W., Martinez-Torres,D., Rahbe,Y., Sabater-Munoz,B., Stern,D., Tagu,D. and Wincker,P.		
TITLE	An expressed sequence tags database for the pea aphid Acyrthosiphon pisum		
JOURNAL	Unpublished (2004)		
COMMENT	Contact: D. Tagu		
INRA Rennes			
UMR BIO3P	BP 35327, F-35653 Le Rheu Cedex France		
Tel:	+33.2.23.48.51.65		
Fax:	+33.2.23.48.51.50		
Risk of contamination by bacterial sequences from obligatory (Buchnera) or facultative endosymbionts. These sequences were obtained in the frame of the International Consortium of Aphid Genomics in collaboration with Genoscope			
PCR Primers	FORWARD: CAGGAACAGCTATGACC		
Plate:	12 row: G column: 12.		
FEATURES			
source	1. 625		
	Location/Qualifiers		
	/organism="Acyrthosiphon pisum"		
	/mol_type="mRNA"		
	/cultivar="developmentstage"		
	/db_xref="taxon:7029"		
	/clone="ID0AAA12CG12"		
	/issue_type="whole insect"		
	/dev_stage="nymphs and adults (parthenogenetic females)"		
	/lab_host="XLI-Blue"		
	/clone_lib="ApMS"		
	/note="Vector: PBS-SK minus; Site 1: EcoRI; Site 2: XhoI; Sample name: ID0AAA ; Plant growth place: Department of Ecology & Evolutionary Biology, Princeton University ; Soil conditions: Soil ; Sowing date: 01/06/1999 ; Harvesting date: 01/06/1999 ; Stress date: no stress ; Description: Aphids inoculated on one-week old Vicia faba under non-sterile conditions. All parthenogenetic stages and both winged and wingless adults were collected for library construction. ; experimental condition: long photoperiod (16-hr light/8-hr dark at 18 c)"		
ORIGIN			
Alignment Scores:			
Pred. No.:	195	Length:	625
Score:	45.00	Matches:	6
Percent Similarity:	84.62%	Conservative:	5
Best Local Similarity:	46.15%	Mismatches:	2
Query Match:	60.81%	Indels:	0
DB:	7	Gaps:	0
US-10-787-382-20 (1-15) x CN754327 (1-625)			
QY	1	PheAlaValGluAsnProMetAsnArgLeuValAlaGlu	13
DB	534	TTTATCATCGAATACCCCACTCAATAGAAATTATCATAGAA	496
RESULT 12			
CN835366/c			
LOCUS	CN835366	902 bp	linear
DEFINITION	ID0ACC28BA10RM1 ID0ACC Acyrthosiphon pisum cDNA clone ID0ACC28BA10		EST 17-NOV-2004
ACCESSION	CN835366		
VERSION	CN835366.1	GI:55800343	

```

VERSION CV835366.1 GI:55801049
KEYWORDS EST.
SOURCE Acyrthosiphon pisum (pea aphid)
ORGANISM Acyrthosiphon pisum
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidiformes;
Aphidoidea; Aphididae; Macrosiphini; Acyrthosiphon.
REFERENCE 1 (bases 1 to 902)
AUTHORS Sabater-Munoz,B., Legeai,F., Bonhomme,J., Dang,P., Dossat,C.,
Duclet,A., Gauthier,J.P., Hunter,W., Martinez-Torres,D., Moya,A.,
Nakabachi,A., Prunier-Leterme,N., Rahbe,Y., Shigenobu,S.,
Simon,J.C., Stern,D., Wincker,P. and Tagu,D.
TITLE Annotated ESTs of the pea aphid
JOURNAL Unpublished (2004)
COMMENT Contact: D. Tagu
INRA Rennes
UMR BIO3P, BP 35327, F-35653 Le Rheu Cedex France
Tel: +33.2.23.48.51.65
Fax: +33.2.23.48.51.50
PCR Primers
FORWARD: CAGGAACAGCTATGACC
Plate: 28B row: A column: 10.
FEATURES
source
1..902
Location/Qualifiers
/organism="Acyrthosiphon pisum"
/mol_type="rRNA"
/cultivar="P123"
/db_xref="taxon:7029"
/clone="ID0ACC28BA10"
/tissue_type="head"
/dev_stage="larvae L3 (parthenogenetic females)"
/lab_host="XLI-Blue"
/clone_lib="ID0ACC"
/note="Vector: pBS-SKminus; Site 1: EcoRI; Site 2: XhoI;
Sample name: ID0ACC ; Plant growth place: INRA Rennes, UMR
BIO3P, 35327, 35653 Le Rheu Cedex France ; Soil
conditions: Soil ; Sowing date: 01/10/2003 ; Harvesting
date: 17/10/2003 ; Description: aphids inoculated on
one-week old Vicia faba germinations under non sterile
conditions experimental condition: long photoperiod (16-hr
light/8-hr dark at 18 degC)"

ORIGIN
Alignment Scores:
Pred. No.: 297 Length: 902
Score: 45.00 Matches: 6
Percent Similarity: 84.62% Conservative: 5
Best Local Similarity: 46.15% Mismatches: 2
Query Match: 60.81% Indels: 0
DB: 8 Gaps: 0

US-10-787-382-20 (1-15) x CV835366 (1-902)

Qy 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGlu 13
|||||
Db 534 TTATCATCGATACCCACCTCAATAGATATCATAGAA 496

RESULT 13
CG458219/c 959 bp DNA linear GSS 17-SBP-2003
LOCUS PUFV93TBC ZM_0.6.1.0 KB Zea mays genomic clone ZMBT0757P18,
genomic survey sequence.
ACCESSION CG458219
VERSION CG458219.1 GI:34843219
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 959)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Reenick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and

Bennetzen,J.
Maize Genomics Consortium
Unpublished (2003)
Other_GSSs: PUFV93TBC
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
FEATURES
Location/Qualifiers
1..959
/organism="Zea mays"
/mol_type="genomic DNA"
/db_xref="taxon:4577"
/clone="ZMBT0757P18"
/clone_lib="ZM_0.6.1.0 KB"
/note="Vector: pCE4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
Cor selected genomic DNA library"

ORIGIN
Alignment Scores:
Pred. No.: 318 Length: 959
Score: 45.00 Matches: 9
Percent Similarity: 78.57% Conservative: 2
Best Local Similarity: 64.29% Mismatches: 3
Query Match: 60.81% Indels: 0
DB: 10 Gaps: 0

US-10-787-382-20 (1-15) x CG458219 (1-959)

Qy 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThr 14
|||||
Db 592 TTCTCTGTGGAGAACCTTAAAGTCGCTCCCGAGAGACC 551

RESULT 14
CA584198 1000 bp mRNA linear EST 19-NOV-2002
LOCUS EST003873 Mycelium and yeast cells from Paracoccidioides
brasiliensis Paracoccidioides brasiliensis cDNA, mRNA sequence.
ACCESSION CA584198.1 GI:25132589
VERSION CA584198.1
KEYWORDS EST.
SOURCE Paracoccidioides brasiliensis
ORGANISM Paracoccidioides brasiliensis
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Onygenales; mitosporic Onygenales; Paracoccidioides.
REFERENCE 1 (bases 1 to 1000)
AUTHORS Felipe,M.S.S., Andrade,R.V., Petrofeza,S.S., Maranhao,A.Q.,
Torres,F.A., Albuquerque,P., Arraes,F.B.M., Arruda,M.,
Azevedo,M.O., Baptista,A.J., Borges,L.A., Borges,C.L., Campos,E.G.,
Cruz,M.R., Daher,B.S., Dantas,A., Ferreira,M.A.S.V., Ghil,G.V.,
Jesuino,R.S.A., Kyaw,C.M., Leitao,L., Martins,C.R., Moraes,L.M.P.,
Neves,E.O., Nicola,A., Alves,E.S., Parente,J.A., Pereira,M.,
Pocas-Fonseca,M.J., Resende,R., Ribeiro,B.M., Saldanha,R.R.,
Santos,S.C., Silva-Pereira,I., Silva,M.A.S., Silveira,E.,
Simoes,I.C., Soares,R.B.A., Souza,D.P., de-Souza,M.T.,
Andrade,E.V., Xavier,M.A.S., Veiga,H.P., Venancio,E.J.,
Carvalho,M.J.A., Oliveira,A.G., Inoue,M.K., Almeida,N.P.,
Walter,M.E.M.T., Soares,C.M.A. and Brígido,M.M.
Transcriptome characterization of the dimorphic and pathogenic
fungus Paracoccidioides brasiliensis by EST analysis
Yeast 20 (3), 263-271 (2003)
12557278
Contact: Felipe MSS
Laboratory of Molecular Biology
Institute of Biology - University of Brasilia
Campus Universitario, Asa Norte, Brasilia, DF 70910-900, BRA
Tel: 55 61 307 2423
Fax: 55 61 349 8411

```

Email: msueli@unb.br
 Seq primer: T7 Sequencing primer.
 Location/Qualifiers
 1..1000
 /organism="Paracoccidioides brasiliensis"
 /mol_type="rRNA"
 /strain="Pb01"
 /db_xref="taxon:121759"
 /clone_lib="Mycelium and yeast cells from Paracoccidioides brasiliensis"
 /note="Pb Lambda Zap Express Library"

ORIGIN

Alignment Scores:
 Pred. No.: 334 Length: 1000
 Score: 45.00 Matches: 8
 Percent Similarity: 78.57% Conservatve: 3
 Best Local Similarity: 57.14% Mismatches: 3
 Query Match: 60.81% Indels: 0
 DB: 6 Gaps: 0

US-10-787-382-20 (1-15) x CA584198 (1-1000)

QY 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThr 14
 Db 723 TTTGTTGTAGACACCCNATAATAGATTAAACGAAACA 764

RESULT 15
 B1084264
 LOCUS
 DEFINITION
 mRNA sequence. 1012 bp mRNA linear EST 20-JUN-2001
 B1084264
 VERSION
 B1084264.1 GI:14502594
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 1 (bases 1 to 1012)
 NIH-MGC http://mgi.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: CLONETECH Laboratories, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1CM1821 row: h column: 18
 High quality sequence stop: 1.
 Location/Qualifiers
 1..1012
 /organism="Homo sapiens"
 /mol_type="rRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5014337"
 /tissue_type="epidermoid carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_102"
 /note="Organ: salivary gland; Vector: pOTB7; Site: 1: XhoI;
 Site: 2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Library constructed
 by Ling Hong in the laboratory of Gerald M. Rubin
 (University of California, Berkeley) using ZAP-cDNA
 synthesis kit (Stratagene), and Superscript II RT (Life
 Technologies). Note: this is a NIH_MGC Library."

FEATURES
 source
 1..1012
 /organism="Lycopersicon esculentum"
 /mol_type="rRNA"
 /cultivar="Micro-Tom"
 /db_xref="taxon:4081"
 /clone="FA35DB12"
 /tissue_type="maturing fruit"
 /clone_lib="Lycopersicon esculentum maturing fruit"

ORIGIN

Alignment Scores:
 Pred. No.: 224 Length: 471
 Score: 44.00 Matches: 8
 Percent Similarity: 76.92% Conservatve: 2
 Best Local Similarity: 61.54% Mismatches: 3
 Query Match: 59.46% Indels: 0
 DB: 3 Gaps: 0

US-10-787-382-20 (1-15) x BP886906 (1-471)

QY 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGlu 13
 Db 13 TTTCTTGTGTAACCCCACTTAATAGATTGATAAAGAA 51

ORIGIN

RESULT 17
 BM281788/c

Alignment Scores:
 Pred. No.: 338 Length: 1012
 Score: 45.00 Matches: 9
 Percent Similarity: 81.82% Conservatve: 0
 Best Local Similarity: 81.82% Mismatches: 2
 Query Match: 60.81% Indels: 0
 DB: 2 Gaps: 0

US-10-787-382-20 (1-15) x B1084264 (1-1012)

QY 3 ValGluAsnProMetAsnArgLeuValAlaGlu 13
 Db 859 GTGCGAAACCCGATGATAGACTATGCGCAGAA 891

RESULT 16
 BP886906
 LOCUS
 DEFINITION
 mRNA sequence. 471 bp mRNA linear EST 16-MAR-2005
 BP886906
 VERSION
 BP886906.1 GI:58230377
 KEYWORDS
 EST.
 SOURCE
 Lycopersicon esculentum (Solanum lycopersicum)
 ORGANISM
 Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
 1 (bases 1 to 471)
 Yamamoto, N., Tsugane, T., Watanabe, M., Yano, K., Maeda, F., Kuwata, C.,
 Moez, T., Nishimura, S. and Shibata, D.
 Expressed sequence tags from the laboratory-grown miniature tomato
 (Lycopersicon esculentum) cultivar Micro-Tom and mining for single
 nucleotide polymorphisms and insertions/deletions in tomato
 cultivars
 Unpublished (2005)
 Contact: Daisuke Shibata
 Kazusa DNA Research Institute;
 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
 Tel: 81-438-52-3947
 Fax: 81-438-52-3948
 Email: shibata@kazusa.or.jp
 This clone was obtained at our laboratory.
 Please visit our web site
 URL: http://www.kazusa.or.jp/jsol/microtom/indexj.html (in Japanese)
 URL: http://www.kazusa.or.jp/jsol/microtom/indexj.html (in English).

FEATURES

source
 1..471
 /organism="Lycopersicon esculentum"
 /mol_type="rRNA"
 /cultivar="Micro-Tom"
 /db_xref="taxon:4081"
 /clone="FA35DB12"
 /tissue_type="maturing fruit"
 /clone_lib="Lycopersicon esculentum maturing fruit"

ORIGIN

Alignment Scores:
 Pred. No.: 224 Length: 471
 Score: 44.00 Matches: 8
 Percent Similarity: 76.92% Conservatve: 2
 Best Local Similarity: 61.54% Mismatches: 3
 Query Match: 59.46% Indels: 0
 DB: 3 Gaps: 0

US-10-787-382-20 (1-15) x BP886906 (1-471)

QY 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGlu 13
 Db 13 TTTCTTGTGTAACCCCACTTAATAGATTGATAAAGAA 51

RESULT 17
 BM281788/c

Fax: +49 6221 4038 290
Email: Thomas.Schluter@lionbioscience.com.

FEATURES

source

```
1. .572
Location/Qualifiers
/organism="Canis familiaris"
/mol_type="mRNA"
/strain="Beagle"
/db_xref="taxon:9615"
/tissue_type="liver"
/dev_stages="adult"
/lab_host="DH10B"
/clone_lib="DG42-liver"
/note="Organ: liver; Vector: Dog pBluescript LION"
```

ORIGIN

Alignment Scores:
Pred. No.: 280 Length: 572
Score: 44.00 Matches: 9
Percent Similarity: 78.57% Conservativeness: 2
Best Local Similarity: 64.29% Mismatches: 3
Query Match: 59.46% Indels: 0
DB: 7 Gaps: 0

US-10-787-382-20 (1-15) x C0675844 (1-572)

QY 2 AlavalGluAenPrometAsnArgLeuValalagluThrLeu 15
|||||
Db 439 GCTTTGAGAACCCCATGGCGAGCTTCGATCCGAACCTCG 480

RESULT 20

AG132831

LOCUS

DEFINITION AG132831 675 bp DNA linear GSS 04-NOV-2001

ACCESSION AG132831 BAC end sequences of Library PTB

VERSION AG132831.1 GI:16662509

KEYWORDS GSS.

SOURCE Pan troglodytes (chimpanzee)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Pan.

REFERENCE

AUTHORS

1 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,

Totoki, Y., Watanabe, H. and Sakaki, Y.

BAC end sequences of Library PTB

Unpublished

2 (bases 1 to 675)

Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,

Totoki, Y., Watanabe, H. and Sakaki, Y.

Direct Submission

Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical

and Chemical Research (RIKEN), Genomic Sciences Center (GSC);

1-7-22 Suehiro-chou, Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan

(E-mail: chimpanzee@riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,

Tel: 81-45-503-9111, Fax: 81-45-503-9170)

Clones are derived from the chimpanzee BAC library PTB This BAC end

was generated during the R&D process and may have higher chance of

clone tracking errors.

PRIMERS

Sequencing: -21M13

LIBRARY

Vector : pKS145

R.Site 1 : SacI

R.Site 2 : SacI

Location/Qualifiers

1. .675

/organism="Pan troglodytes"

/mol_type="genomic DNA"

/db_xref="taxon:9598"

/clone="PTB-145F04.F"

/sex="male"

/cell_type="lymphoblast"

/clone_lib="PTB Chimpanzee Male BAC Library"

FEATURES

source

ORIGIN

Alignment Scores:

Pred. No.: 339 Length: 675
Score: 44.00 Matches: 7
Percent Similarity: 78.57% Conservativeness: 4
Best Local Similarity: 59.00% Mismatches: 3
Query Match: 59.46% Indels: 0
DB: 10 Gaps: 0

US-10-787-382-20 (1-15) x AG132831 (1-675)

QY 2 AlavalGluAenPrometAsnArgLeuValalagluThrLeu 15
|||||
Db 381 GCAATTCAGAACCTCAGAACGACATCTCAGAACATTG 422

RESULT 21

BI731326/C

LOCUS

DEFINITION BI731326 797 bp mRNA linear EST 20-SEP-2001

ACCESSION BI731326

VERSION BI731326.1 GI:15708339

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
1 (bases 1 to 797)

REFERENCE

AUTHORS

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: The Cepko Laboratory

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM11917 row: b column: 14

High quality sequence stop: 605.

FEATURES

source

1. .797

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="IMAGE:5360557"

/tissue_type="retina"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_94"

/note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;

Site 2: SalI; Cloned unidirectionally; oligo-dT primed.

Average insert size 3.3 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.: 409 Length: 797
Score: 44.00 Matches: 8
Percent Similarity: 78.57% Conservativeness: 3
Best Local Similarity: 57.14% Mismatches: 3
Query Match: 59.46% Indels: 0
DB: 3 Gaps: 0

US-10-787-382-20 (1-15) x BI731326 (1-797)

QY 2 AlavalGluAenPrometAsnArgLeuValalagluThrLeu 15
|||||
Db 139 GCTGTGAGAACCCCATAGCTGCTGGTGGGAGAACATC 98

```

RESULT 22
CL670791
LOCUS
DEFINITION
    CL670791 817 bp DNA linear GSS 09-JUL-2004
    PRI0163a.C12 - PRI0163a.B21 (817) Mixed stage fosmid library of P.
    pacificus var. California Pristionchus pacificus genomic, genomic
    survey sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
    Pristionchus pacificus
    Pristionchus pacificus
    Pristionchus pacificus
    Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
    Neodiplogasteridae; Pristionchus.
REFERENCE
AUTHORS
    Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.
TITLE
    AppADB: an AcedB database for the nematode satellite organism
JOURNAL
Nucleic Acids Res. 32 (1), D421-D422 (2004)
PUBMED
14681447
COMMENT
    Contact: Sommer RJ
    Evolutionary Biology
    Max-Planck-Institute for Developmental Biology
    Spemannstr. 37-39, Tuebingen D-72076, Germany
    Tel: 00497071601371
    Fax: 00497071601498
    Email: ralf.sommer@tuebingen.mpg.de
    This library was generated at Caltech, Pasadena, USA and end
    sequenced at Vancouver, Canada.
    Seq primer: T7
    Class: fosmid ends.
    Location/Qualifiers
        1..817
        /organism="Pristionchus pacificus"
        /mol_type="genomic DNA"
        /strain="California"
        /db_xref="taxon:54126"
        /clone_lib="Mixed stage fosmid library of P. pacificus
        var. California"
        /note="Vector: pEpifos-5 Fosmid vector"
FEATURES
    source
        1..817
        /organism="Pristionchus pacificus"
        /mol_type="genomic DNA"
        /strain="California"
        /db_xref="taxon:54126"
        /clone_lib="Mixed stage fosmid library of P. pacificus
        var. California"
        /note="Vector: pEpifos-5 Fosmid vector"
ORIGIN
    Alignment Scores:
    Pred. No.: 421 Length: 817
    Score: 44.00 Matches: 8
    Percent Similarity: 76.92% Conservative: 2
    Best Local Similarity: 61.54% Mismatches: 3
    Query Match: 59.46% Indels: 0
    DB: 10 Gaps: 0
    US-10-787-382-20 (1-15) x CL670791 (1-817)
    QY 2 AlavalGluAsnProMetAsnArgLeuValAlaGluThr 14
    |||::: ||||| ||| |||||:::
    Db 326 GCGATAACACCCGCGAARACGCGCTCGCTCAAC 364

RESULT 23
CL665378/c
LOCUS
DEFINITION
    CL665378 832 bp DNA linear GSS 09-JUL-2004
    PRI0149c.C09 - PRI0149c.B21 (832) Mixed stage fosmid library of P.
    pacificus var. California Pristionchus pacificus genomic, genomic
    survey sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
    Pristionchus pacificus
    Pristionchus pacificus
    Pristionchus pacificus
    Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
    Neodiplogasteridae; Pristionchus.
REFERENCE
AUTHORS
    Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.
TITLE
    AppADB: an AcedB database for the nematode satellite organism
JOURNAL
Nucleic Acids Res. 32 (1), D421-D422 (2004)
PUBMED
14681447
COMMENT
    Contact: Sommer RJ
    Evolutionary Biology
    Max-Planck-Institute for Developmental Biology
    Spemannstr. 37-39, Tuebingen D-72076, Germany
    Tel: 00497071601371
    Fax: 00497071601498
    Email: ralf.sommer@tuebingen.mpg.de
    This library was generated at Caltech, Pasadena, USA and end
    sequenced at Vancouver, Canada.
    Seq primer: T7
    Class: fosmid ends.
    Location/Qualifiers
        1..832
        /organism="Pristionchus pacificus"
        /mol_type="genomic DNA"
        /strain="California"
        /db_xref="taxon:54126"
        /clone_lib="Mixed stage fosmid library of P. pacificus
        var. California"
        /note="Vector: pEpifos-5 Fosmid vector"
ORIGIN
    Alignment Scores:
    Pred. No.: 421 Length: 832
    Score: 44.00 Matches: 8
    Percent Similarity: 76.92% Conservative: 2
    Best Local Similarity: 61.54% Mismatches: 3
    Query Match: 59.46% Indels: 0
    DB: 10 Gaps: 0
    US-10-787-382-20 (1-15) x CL665378 (1-832)
    QY 2 AlavalGluAsnProMetAsnArgLeuValAlaGluThr 14
    |||::: ||||| ||| |||||:::
    Db 259 GCGATAACACCCGCGAARACGCGCTCGCTCAAC 221

RESULT 24
BF977950
LOCUS
DEFINITION
    BF977950 966 bp mRNA linear EST 22-JAN-2001
    602148422F1 NIH_MGC_62 Homo sapiens cDNA clone IMAGE:4307453 5',
    mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
    Homo sapiens (human)
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
    Homnidae; Homo.
REFERENCE
    1 (bases 1 to 966)
    NIH-MGC http://mgc.nci.nih.gov/.
    National Institutes of Health, Mammalian Gene Collection (MGC)
    Unpublished (1999)
    Contact: Robert Strausberg, Ph.D.
    Email: cgabbs-r@mail.nih.gov
    Tissue Procurement: ATCC/DCTD/BTP
    cDNA library preparation: CLONETECH Laboratories, Inc.
    cDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
    DNA Sequencing by: Incyte Genomics, Inc.
    Clone distribution: MGC clone distribution information can be
    found through the I.M.A.G.E. Consortium/LLNL at:
    http://image.llnl.gov
    Plate: LLCM1177 row: k column: 06
    High quality sequence stop: 84.
    Location/Qualifiers
        1..966
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:4307453"
        /tissue_type="melanotic melanoma, high MDR"

```



```

/lab_host="PH10B (T1 phage-resistant)"
/clone_lib="NIH MGC 62"
/note="Organ: skin; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggccgctggcc); Site_2: SfiI (ggcattatggcc);
Double-stranded cDNA was prepared from cell line RNA. 5'
and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CAGGGCATTATGGCC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCGGAGCGGCACATG-dt(30)BN-3'
(where B = A, C, G or T). Average
insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA)."

ORIGIN
Alignment Scores:
Pred. No.: 510 Length: 966
Score: 44.00 Matches: 8
Percent Similarity: 90.91% Conservative: 2
Best Local Similarity: 72.73% Mismatches: 1
Query Match: 59.46% Indels: 0
DB: 2 Gaps: 0

US-10-787-382-20 (1-15) x BF977950 (1-966)

Qy 1 PheAlaValGluAsnProMetAsnArgLeuVal 11
|||||
Db 859 TTTGCTGTTGAAATCTCTCAGCGGTTTATTA 891
|||||

RESULT 25
CNS0764C 1011 bp DNA linear GSS 07-JUL-2001
LOCUS clone XBA0AB001G09 of library XBA0AB from strain CLIB 210 of
DEFINITION Kluyveromyces lactis, genomic survey sequence.
ACCESSION AL430930
VERSION AL430930.1 GI:12214124
KEYWORDS GSS.
SOURCE Kluyveromyces lactis
ORGANISM Kluyveromyces lactis
REFERENCE Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
AUTHORS Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
1 (bases 1 to 1011)
Bouciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
Solciet,Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
Malbertuy,A., Neuvelise,C., Ozier-Kalogeropoulos,O., Potier,S.,
Saurin,W., Tekai,P., Toffano-Nioche,C., Wesolowski-Louvel,M.,
Wincker,P. and Weissenbach,J.
Genomic exploration of the hemiascomycetous yeasts: 1. A set of
yeast species for molecular evolution studies
FEBS Lett. 487 (1), 3-12 (2000)
11152876
2 (bases 1 to 1011)
Bolotin-Fukuhara,M., Toffano-Nioche,C., Artiguenave,F.,
Duchateau-Nguyen,G., Lemaire,M., Marmelise,R., Montrocher,R.,
Robert,C., Termier,M., Wincker,P. and Wesolowski-Louvel,M.
Genomic exploration of the hemiascomycetous yeasts: 11.
Kluyveromyces lactis
FEBS Lett. 487 (1), 56-70 (2000)
11152886
3 (bases 1 to 1011)
Genoscope.
Direct Submission
Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,
2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
seqrefgenoscope.cns.fr - Web : www.genoscope.cns.fr)
This GSS is part of a random genomic sequencing program of thirteen
yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
5 kb were prepared and both extremities were sequenced. See
keywords for description of this sequence and for the sequence of
the other extremity of this insert.

FEATURES
source
1..1011
Location/Qualifiers
/organism="Kluyveromyces lactis"
/mol_type="genomic DNA"
/strain="CLIB 210"
/variety="lactis"
/db_xref="taxon:28985"
/clone="XBA0AB001G09"
/clone_lib="XBA0AB"
<148..>465
/note="similar to Saccharomyces cerevisiae ORF YCR081w [
SRB8 ; DNA-directed RNA polymerase II holoenzyme and Srb10
CDK subcomplex subunit 1]"
/evidence=not_experimental
complement(<548..>1003)
/note="similar to Saccharomyces cerevisiae ORF YNR039c [
weak similarity to Anopheles mitochondrial NADH
dehydrogenase subunit 2]"
/evidence=not_experimental

misc_feature
537 Length: 1011
Pred. No.: 44.00 Matches: 8
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 80.00% Mismatches: 0
Query Match: 59.46% Indels: 0
DB: 11 Gaps: 0

US-10-787-382-20 (1-15) x CNS0764C (1-1011)

Qy 2 AlaValGluAsnProMetAsnArgLeuVal 11
|||||
Db 948 GCAGTAGAAATCCCATTAATAGATTGTT 977
|||||

RESULT 26
CNS05P3H/C 1016 bp DNA linear GSS 01-SEP-2000
LOCUS Tetraodon nigroviridis genome survey sequence SP6 end of clone
DEFINITION 038M15 of library B from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION AL347462
VERSION AL347462.1 GI:8241232
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontoidea; Tetraodontidae; Tetraodon.
1
Roest Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
Saurin,W. and Weissenbach,J.
Estimate of human gene number provided by genome-wide analysis
using Tetraodon nigroviridis DNA sequence
Nat. Genet. 25 (2), 235-238 (2000)
10835645
2
Roest Crolius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C.,
Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
Saurin,W., Bernot,A. and Weissenbach,J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
Genome Res. 10 (7), 939-949 (2000)
10899143
3 (bases 1 to 1016)
Genoscope.
Direct Submission
Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :

```


TITLE
JOURNAL
COMMENT
 Brokstein, P. and Lindquist, E.A.
 DOE Joint Genome Institute Xenopus tropicalis EST project
 Unpublished (2004)
 Other_ESTS: JGI_XZT42776.fwd
 Contact: Lindquist, E.A., Richardson, P.
 DOE Joint Genome Institute
 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 Tel: 925 296 5600
 Fax: 925 296 5710
 Email: cdna@jgi-psf.org
 Tissue Procurement: Richard M. Harland Laboratory, University of California, Berkeley: http://tropicalis.berkeley.edu/home
 cDNA Library Preparation: Richard M. Harland Laboratory, University of California, Berkeley
 DNA Sequencing: DOE Joint Genome Institute: http://www.jgi.doe.gov
 Clone Distribution: I.M.A.G.E. Consortium/LiNL: http://image.llnl.gov
 Naming Conventions: EST name is generated by the concatenation of the JGI Clone ID and the direction of sequencing. The suffix '.rev' indicates a reverse sequencing read of the insert. It does not necessarily reflect the orientation of the insert.
 Poly-A: Based upon the presence of a run of 14 or more T residues at the beginning of the sequence, this clone was polyadenylated.
 The resulting Poly-T sequence has been removed.
 Small Insert: Based upon one or more sequencing reads of this clone where vector sequence was present at both ends, this clone has been determined to contain a cDNA insert on the order of 600-1000 bases.
 Plate: XZT 0445 row: P column: 13
 High quality sequence stop: 263
 POLYA=Yes.

FEATURES **Source**

1..494
 Location/Qualifiers
 /organism="Xenopus tropicalis"
 /mol_type="mRNA"
 /db_xref="taxon:8364"
 /clone="IMAGE:7618431"
 /tissue_type="whole embryo"
 /dev_stage="Tadpole (st. 36-41)"
 /lab_host="E. coli XLI-Blue derivative, Stratagene Electorfen-Blue"
 /clone_lib="NIH XGC tropTad5"
 /notes="Vector: pCS108; Site 1: SalI; Site 2: NotI; Tadpole library constructed by Russell B. Fletcher in R. Harland's lab using poly A RNA and oligo dT primers (Invitrogen SuperScript Plasmid System for cDNA Synthesis and Cloning). SalI (5' end) -NotI (3' end) cDNA was inserted into vector pCS108 (http://mcb.berkeley.edu/labs/harland/pages/plasmids.html)".

ORIGIN

Alignment Scores:
 Pred. No.: 377 Length: 494
 Score: 43.00 Matches: 10
 Percent Similarity: 66.67% Conservative: 0
 Best Local Similarity: 66.67% Mismatches: 5
 Query Match: 58.11% Indels: 0
 DB: 8 Gaps: 0

US-10-787-382-20 (1-15) x DR897668 (1-494)

QY 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
Db 205 TTTCGGTTCCTCAATCAATGAATTGTTGTGGCCGCCGTTTG 161

RESULT 32
CE797827/c
LOCUS
DEFINITION
 tigr-gss-dog-17000317573763 Dog Library Canis familiaris genomic, genomic survey sequence.
ACCESSION
VERSION
KEYWORDS
 CE797827 GI:37138596
 GSS.

SOURCE **ORGANISM**

Canis familiaris (dog)
 Canis familiaris
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
 Canis.

REFERENCE **AUTHORS**

1 (bases 1 to 511)
 Kirkness, E.F., Bafna, V., Halpern, A.L., Levy, S., Remington, K.,
 Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M. and
 Venter, J.C.
 The dog genome: survey sequencing and comparative analysis
 Science 301 (5641), 1898-1903 (2003)

TITLE **JOURNAL** **PUBMED** **COMMENT**

Contact: Kirkness EF
 The Institute for Genomic Research
 Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
 Rockville, MD 20850, USA
 Tel: 301-838-0200
 Fax: 301-838-0208
 Email: ekirknes@tigr.org
 Class: shotgun.

FEATURES **source**

1..511
 Location/Qualifiers
 /organism="Canis familiaris"
 /mol_type="genomic DNA"
 /strain="Standard Poodle"
 /db_xref="taxon:9615"
 /clone_lib="Dog Library"
 /note="Site 1: BstXI; Libraries were prepared from peripheral blood"

ORIGIN

Alignment Scores:
 Pred. No.: 392 Length: 511
 Score: 43.00 Matches: 7
 Percent Similarity: 80.00% Conservative: 5
 Best Local Similarity: 46.67% Mismatches: 3
 Query Match: 58.11% Indels: 0
 DB: 10 Gaps: 0

US-10-787-382-20 (1-15) x CE797827 (1-511)

QY 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
Db 113 TTTCAGTAGAATAATCCACTTTAGAAATAATTCAAATACCTTA 69

RESULT 33

CC062239
LOCUS
DEFINITION
 ugma001e002a11f1 ugma unfiltered library (LibID: 146) Glycine max genomic clone ugma001e002a11, genomic survey sequence.

ACCESSION **VERSION** **KEYWORDS** **SOURCE** **ORGANISM**

CC062239.1 GI:29841621
 GSS.
 Glycine max (soybean)
 Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.

REFERENCE **AUTHORS**

1 (bases 1 to 517)
 Budiman, M.A., Freese, R.G., Flick, E., Citek, R.W., Nunberg, A.,
 Robbins, D., Bedell, J.A. and Lakey, N.

TITLE **JOURNAL** **COMMENT**

Genethresher methylation filtered genomic sequences from soybean
 Unpublished (2003)
 Contact: Bedell JA
 Orion Genomics, LLC
 4041 Forest Park Ave, St. Louis, MO 63108, USA
 Tel: 314 615 6979
 Fax: 314 615 5975
 Email: jbedell@oriongenomics.com
 Plate: ugma001e002 row: a column: 11
 Seq primer: f Forward
 Class: shotgun

Sali/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 3.8 Kb. The library was constructed by Yulan Piao."

ORIGIN

Alignment Scores:
Pred. No.: 414 Length: 536
Score: 43.00 Matches: 8
Percent Similarity: 100.00% Conservative: 4
Best Local Similarity: 66.67% Mismatches: 0
Query Match: 58.11% Indels: 0
DB: 6 Gaps: 0

US-10-787-382-20 (1-15) x CA877602 (1-536)

Qy 1 PheAlaValGluAsnProMetAsnArgLeuValAla 12
Db 364 TGGGCCCTTGAGATCCTATTTCAGAGATTAGTTGCC 399

RESULT 36

CC578117

LOCUS CH240_456D24.TARBAC13P2 CHORI-240 Bos taurus genomic clone
DEFINITION CH240_456D24, genomic survey sequence.

ACCESSION CC578117.1 GI:31922460

VERSION GSS.

KEYWORDS Bos taurus (cow)

SOURCE Bos taurus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Bovinae; Bos.

REFERENCE

AUTHORS

Holt, R., Stott, J., Yang, G., Barber, S., Smalish, D., Prabhu, A.-L.,
Tsai, M., Cloutier, A., Lee, D., Ginn, N., Olson, T., Mayo, M.,
Butterfield, Y., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R.,
Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S.,
Schein, J., Marra, M., de Jong, P., Keele, J.W. and Kappes, S.M.
Bovine BAC End Sequences from Library CHORI-240, PLATES 399 to 478
Unpublished (2003)

TITLE

Other GSSs: CH240_456D24.T7

JOURNAL

COMMENT

Contact: Rob Holt
Sequencing
The British Columbia Cancer Agency Genome Science Centre
600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4B6
Tel: 604-877-6085
Fax: 604-877-6276
Email: rholt@bcgsc.ca

Clones are derived from the bovine BAC library CHORI-240
(<http://www.chori.org/bacpac/bovine240.htm>). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(<http://www.chori.org/bacpac/ordering/information.htm>). This work
was undertaken as part of the International Bovine BAC Mapping
Consortium (IBMC) by CSIRO Livestock Industries, Australia and the
British Columbia Genome Sciences Centre, Canada.
Plate: 456 row: D column: 24
Seq primer: SP6
Class: BAC ends.

FEATURES

source

Location/Qualifiers
1..553
/organism="Bos taurus"
/mol_type="genomic DNA"
/strain="Breed: Hereford"
/db_xref="taxon:9913"
/clone="CH240_456D24"
/sex="Male"
/cell_type="Blood"
/clone_lib="CHORI-240"
/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC
library (Male) produced by Pieter de Jong"

ORIGIN

Alignment Scores:
Pred. No.: 429 Length: 553
Score: 43.00 Matches: 8
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 88.89% Mismatches: 0
Query Match: 58.11% Indels: 0
DB: 9 Gaps: 0

US-10-787-382-20 (1-15) x CC578117 (1-553)

Qy 1 PheAlaValGluAsnProMetAsnArg 9
Db 374 TTTGCCGTGGAGATCCCATGGACAGA 400

RESULT 37

CG841296

LOCUS Ynhw1910 HW-YUBAC Bos taurus genomic clone linear GSS 01-SEP-2004
DEFINITION 5', genomic survey sequence.

ACCESSION CG841296

VERSION CG841296.1 GI:51786324

KEYWORDS GSS.

SOURCE Bos taurus (cow)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Bovinae; Bos.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Genomic sequences from Korean Cattle (Hanwoo) blood
Unpublished (2003)
Contact: Inho Choi
Molecular Biology
Yeungnam University
214-1, Dae-dong Gyeongsan, Korea 712-749
Tel: 82 53 810 2933
Fax: 82 53 816 3637
Email: inhochoi@yumail.ac.kr
Seq primer: T7 Forward
Class: BAC ends.

FEATURES

source

Location/Qualifiers
1..562
/organism="Bos taurus"
/mol_type="genomic DNA"
/strain="Korean Cattle (Hanwoo)"
/db_xref="taxon:9913"
/clone="HW-YUBAC2-184-1-D03-T7"
/sex="Male"
/tissue_type="Blood"
/cell_type="Leucocyte"
/dev_stage="Adult"
/clone_lib="HW-YUBAC"
/note="Vector: pIndigoBAC-5"

ORIGIN

Alignment Scores:
Pred. No.: 437 Length: 562
Score: 43.00 Matches: 8
Percent Similarity: 81.82% Conservative: 1
Best Local Similarity: 72.73% Mismatches: 2
Query Match: 58.11% Indels: 0
DB: 10 Gaps: 0

US-10-787-382-20 (1-15) x CG841296 (1-562)

Qy 1 PheAlaValGluAsnProMetAsnArgLeuVal 11

Db 184 TTTTCCGTGGAGATCCCATGAACAGAGATC 216

RESULT 38

AZ759493

LOCUS AZ759493 591 bp DNA linear GSS 16-FEB-2001
 DEFINITION 1M0552K03F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 clone UUGC1M0552K03 F, genomic survey sequence.
 ACCESSION AZ759493
 VERSION AZ759493
 KEYWORDS AZ759493.1 GI:12866341
 SOURCE GSS.
 ORGANISM Mus musculus (house mouse)
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridea; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 591)
 AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
 Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,
 Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von
 Niederhausern, A. and Wright, D., Weiss, R.
 TITLE Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 JOURNAL Unpublished (2000)
 COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0552 row: K column: 03
 Seq primer: CGTGTGAACGACGCCAGT
 Class: plasmid ends
 High quality sequence stop: 591.
 Location/Qualifiers
 1..591
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC1M0552K03"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /note="Vector: PWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male); was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptor DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of pWD42 (GI:4732114|gb|AF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adaptorised mouse DNA was annealed to
 adaptor vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

FEATURES

source
 1..591
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC1M0552K03"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /note="Vector: PWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male); was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptor DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of pWD42 (GI:4732114|gb|AF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adaptorised mouse DNA was annealed to
 adaptor vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

ORIGIN

Alignment Scores:
 Pred. No.: 463 Length: 591
 Score: 43.00 Matches: 8
 Percent Similarity: 75.00% Conservative: 1
 Best Local Similarity: 66.67% Mismatches: 3
 Query Match: 58.11% Indels: 0
 DB: 9 Gaps: 0

US-10-787-382-20 (1-15) x AZ759493 (1-591)

Qy
 Db
 RESULT 39
 CF799042
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

CF799042 598 bp mRNA linear EST 21-OCT-2003
 Lr_PAHCF_13B11_M13R Earthworm Fluorantene Exposure Library
 Lumbricus rubellus cDNA clone Lr_PAHCF_13B11 5', mRNA sequence.
 CF799042
 CF799042.1 GI:37803612
 EST.
 Lumbricus rubellus (humus earthworm)
 Lumbricus rubellus
 Eukaryota; Metazoa; Annelida; Clitellata; Oligochaeta; Haplotaxida;
 Lumbricina; Lumbricidae; Lumbricus.
 1 (bases 1 to 598)
 Chaiseley, J., Hedley, B.A., Morgan, J.C., Sturzenbaum, S., Kille, P. and
 Blaxter, M.
 The Lumbricus rubellus EST program - Sequences from a Fluorantene
 Exposure library
 Unpublished (2003)
 Contact: Jennifer Chaiseley
 BIOSI 1
 Cardiff University
 Main College, Museum Avenue, Cardiff, CF11 3TL, UK
 Tel: +44 2920876680
 Fax: +44 2920874305
 Email: chaiseley@cardiff.ac.uk, Kille@cardiff.ac.uk
 Sequencing was performed in Cardiff using the pBluescriptII XR cDNA
 library (Stratagene) protocol.
 PCR Primers
 FORWARD: gtttccagtcacgacg
 BACKWARD: caggaacagctatgaccatg
 Plate: 13 row: B Column: 11
 Seq primer: caggaacagctatgaccatg
 High quality sequence stop: 496.
 Location/Qualifiers
 1..598
 /organism="Lumbricus rubellus"
 /mol_type="mRNA"
 /db_xref="taxon:35632"
 /clone="Lr_PAHCF_13B11"
 /tissue_type="Whole worm"
 /dev_stage="Adult"
 /clone_lib="Earthworm Fluorantene Exposure Library"
 /note="Vector: pBluescript II SK+; The library was
 prepared using protocols given by the supplier
 (Stratagene)."

FEATURES
 source
 1..598
 /organism="Lumbricus rubellus"
 /mol_type="mRNA"
 /db_xref="taxon:35632"
 /clone="Lr_PAHCF_13B11"
 /tissue_type="Whole worm"
 /dev_stage="Adult"
 /clone_lib="Earthworm Fluorantene Exposure Library"
 /note="Vector: pBluescript II SK+; The library was
 prepared using protocols given by the supplier
 (Stratagene)."

ORIGIN

Alignment Scores:
 Pred. No.: 469 Length: 598
 Score: 43.00 Matches: 8
 Percent Similarity: 91.67% Conservative: 3
 Best Local Similarity: 66.67% Mismatches: 1
 Query Match: 58.11% Indels: 0
 DB: 6 Gaps: 0

ORIGIN

US-10-787-382-20 (1-15) x CF799042 (1-598)
 Qy
 Db
 RESULT 40
 CF799042
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS

US-10-787-382-20 (1-15) x CF799042 (1-598)
 Qy
 Db
 RESULT 40
 CF799042
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS

US-10-787-382-20 (1-15) x CF799042 (1-598)
 Qy
 Db
 RESULT 40
 CF799042
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS

SOURCE
ORGANISM
Medicago truncatula (barrel medic)
Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.

REFERENCE
1 (bases 1 to 608)
Town,C.D., Shetty,J., Koo,H. and Feldblyum,T.F.
Sequencing of BAC ends from Medicago truncatula
Unpublished (2003)
JOURNAL
COMMENT
Other GSSs: MBFHF66TR
Contact: Chris Town

TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.

Tel: 301-838-3523

Fax: 301-838-0208

Email: cdtown@tigr.org

Seq primer: TGTAAACGACGGCCAGT

Class: BAC ends.

FEATURES
Location/Qualifiers

1..608
/organism="Medicago truncatula"
/mol_type="genomic DNA"
/cultivar="genotype A17"
/db_xref="taxon:3880"
/clone="43L12"
/clone_lib="meth2"
/note="Vector: pSelOBAC11; Site_1: HindIII; Site_2:
HindIII; Cook, D.R. and Kim, D.J, unpublished"

ORIGIN

Alignment Scores:
Pred. No.: 478 Length: 608
Score: 43.00 Matches: 7
Percent Similarity: 100.00% Conservative: 3
Best Local Similarity: 70.00% Mismatches: 0
Query Match: 58.11% Indels: 0
DB: 10 Gaps: 0

US-10-787-382-20 (1-15) x CG950325 (1-608)

QY 2 AlaValGluAsnProMetAsnArgLeuVal 11

Db 484 GCATTGAAATCCATTGAATGCTCTCATT 455

RESULT 41

CW171674
LOCUS
DEFINITION
613 bp DNA linear GSS 29-OCT-2004
104_582_11155869_148_36530_087 Sorghum methylation filtered library
(LibID: 104) Sorghum bicolor genomic clone 11155869, genomic survey
sequence.

ACCESSION
CW171674

VERSION
GI:54864241

KEYWORDS
GSS.

SOURCE
Sorghum bicolor (sorghum)

ORGANISM

Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.

REFERENCE
1 (bases 1 to 613)

AUTHORS

Bedell,J.A., Budiman,M.A., Nunberg,A., Citek,R.W., Robbins,D.,
Jones,J., Flick,E., Rohlfing,T., Fries,J., Bradford,K.,
McMenamy,J., Smith,M., Holsman,H., Roe,B.A., Wiley,G., Korf,I.F.,
Rabinowicz,P.D., Lakey,N., McCombie,W.R., Jeddeloh,J.A. and
Martensen,R.A.

Sorghum genome sequencing by methylation filtration

PLOS Biol. 3 (1), e13 (2005)

15660154

COMMENT

Contact: Bedell JA
Orion Genomics, LLC
4041 Forest Park Ave, St. Louis, MO 63108, USA
Tel: 314 615 6979
Fax: 314 615 5975

Email: jbedell@oriongenomics.com
Plate: 582 row: j column: 21
Seq primer: SWfor Forward
Class: methylation filtered
High quality sequence stop: 613.
Location/Qualifiers

FEATURES

source

1..613

/organism="Sorghum bicolor"

/mol_type="genomic DNA"

/cultivar="ATX623"

/db_xref="taxon:4558"

/clone="11155869"

/clone_lib="Sorghum methylation filtered library (LibID:
104)"

/note="Organ: leaf; Vector: pBCSK(-); Site_1: HincII; DNA
prepared from purified nuclei was randomly sheared,
end-repaired, size fractionated to enrich for the 0.5 to 5
kb fraction, ligated into HincII-digested pBCSK(-) vector
and electroporated into E. coli cells. This is a
methylation filtered library."

ORIGIN

Alignment Scores:
Pred. No.: 482 Length: 613
Score: 43.00 Matches: 8
Percent Similarity: 90.91% Conservative: 2
Best Local Similarity: 72.73% Mismatches: 1
Query Match: 58.11% Indels: 0
DB: 10 Gaps: 0

US-10-787-382-20 (1-15) x CW171674 (1-613)

QY 5 AsnProMetAsnArgLeuValAlaGluThrLeu 15

Db 297 AACAACTAAACAGGCTCATTGCAGAAACACTT 329

RESULT 42

CW301651/c

LOCUS

DEFINITION

639 bp DNA linear GSS 31-OCT-2004
104_785_11464311_116_36265_056 Sorghum methylation filtered library
(LibID: 104) Sorghum bicolor genomic clone 11464311, genomic survey
sequence.

ACCESSION
CW301651

VERSION
GI:55017839

KEYWORDS
GSS.

SOURCE
Sorghum bicolor (sorghum)

ORGANISM

Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.

REFERENCE
1 (bases 1 to 639)

AUTHORS

Bedell,J.A., Budiman,M.A., Nunberg,A., Citek,R.W., Robbins,D.,
Jones,J., Flick,E., Rohlfing,T., Fries,J., Bradford,K.,
McMenamy,J., Smith,M., Holsman,H., Roe,B.A., Wiley,G., Korf,I.F.,
Rabinowicz,P.D., Lakey,N., McCombie,W.R., Jeddeloh,J.A. and
Martensen,R.A.

Sorghum genome sequencing by methylation filtration

PLOS Biol. 3 (1), e13 (2005)

15660154

COMMENT

Contact: Bedell JA
Orion Genomics, LLC
4041 Forest Park Ave, St. Louis, MO 63108, USA
Tel: 314 615 6979
Fax: 314 615 5975

Email: jbedell@oriongenomics.com

Plate: 785 row: j column: 15

Seq primer: T3 Reverse

Class: methylation filtered

High quality sequence stop: 639.

FEATURES

source

1..639

/organism="Sorghum bicolor"

/mol_type="genomic DNA"


```

/cultivar="ATx623"
/db_xref="taxon:4558"
/clone="11464311"
/clone_lib="Sorghum methylation filtered library (LibID:
104)"
/note="Organ: leaf; Vector: pBCSK(-); Site 1: HincII; DNA
end-repaired, size fractionated to enrich for the 0.5 to 5
kb fraction, ligated into HincII-digested pBCSK(-) vector
and electroporated into E. coli cells. This is a
methylation filtered library."

```

ORIGIN

```

Alignment Scores:
Pred. No.: 506 Length: 639
Score: 43.00 Matches: 8
Percent Similarity: 90.91% Conservative: 2
Best Local Similarity: 72.73% Mismatches: 1
Query Match: 58.11% Indels: 0
DB: 10 Gaps: 0

```

US-10-787-382-20 (1-15) x CW301651 (1-639)

Qy 5 AsnProMetAsnArgLeuValAlaGluThrLeu 15

Db 397 AACAACTAACAGGCTCATTGCAGAAACACTT 365

RESULT 43

```

CW277572/2
LOCUS CW277572.1 665 bp DNA linear GSS 31-OCT-2004
DEFINITION 104 751_11406099_148 35413_004 Sorghum methylation filtered library
(LibID: 104) Sorghum bicolor genomic clone 11406099, genomic survey
sequence.
ACCESSION CW277572
VERSION CW277572.1 GI:54993760
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE 1 (bases 1 to 665)
AUTHORS Bedell,J.A., Budiman,M.A., Nunberg,A., Citek,R.W., Robbins,D.,
Jones,J., Flick,E., Rohlfing,T., Fries,J., Bradford,K.,
McWenamy,J., Smith,M., Holsman,H., Roe,B.A., Wiley,G., Korf,I.F.,
Rabinowicz,P.D., Lakey,N., McCombie,W.R., Jeddloh,J.A. and
Martienssen,R.A.
Sorghum genome sequencing by methylation filtration
PLOS Biol. 3 (1), e13 (2005)
15660154
Contact: Bedell JA
Orion Genomics, LLC
4041 Forest Park Ave, St. Louis, MO 63108, USA
Tel: 314 615 6979
Fax: 314 615 5975
Email: jbedell@oriongenomics.com
Plate: 751 row: m column: 03
Seq primer: SWfor Forward
Class: methylation filtered
High quality sequence stop: 665.
Location/Qualifiers
1..665
/organism="Sorghum bicolor"
/mol_type="genomic DNA"
/cultivar="ATx623"
/db_xref="taxon:4558"
/clone="11406099"
/clone_lib="Sorghum methylation filtered library (LibID:
104)"
/note="Organ: leaf; Vector: pBCSK(-); Site 1: HincII; DNA
end-repaired, size fractionated to enrich for the 0.5 to 5
kb fraction, ligated into HincII-digested pBCSK(-) vector

```

FEATURES

source

and electroporated into E. coli cells. This is a
methylation filtered library."

ORIGIN

```

Alignment Scores:
Pred. No.: 530 Length: 665
Score: 43.00 Matches: 8
Percent Similarity: 90.91% Conservative: 2
Best Local Similarity: 72.73% Mismatches: 1
Query Match: 58.11% Indels: 0
DB: 10 Gaps: 0

```

US-10-787-382-20 (1-15) x CW277572 (1-665)

Qy 5 AsnProMetAsnArgLeuValAlaGluThrLeu 15

Db 540 AACAACTAACAGGCTCATTGCAGAAACACTT 508

RESULT 44

BB303483/3

LOCUS BB303483.2

DEFINITION BB303483 RIKEN full-length enriched, adult male corpora

Quadrigenina Mus musculus cDNA clone B230104L09 3', mRNA sequence.

ACCESSION BB303483

VERSION BB303483.2 GI:16402059

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridea; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 678)

AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,

Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,

Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,

Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,

Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,

Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,

Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayaishizaki,Y.

RIKEN Mouse ESTs (Arakawa,T., et al. 2001)

Unpublished (2001)

TITLE On Jul 10, 2000 this sequence version replaced gi:9004188.

JOURNAL Contact: Yoshihide Hayaishizaki

COMMENT Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Sucho-cho, Tsukuba-shi, Ibaraki, Japan

Tel: 81-45-503-3222

Fax: 81-45-503-3216

Email: genome-res@gsr.riken.jp, URL: http://genome.gsc.riken.jp/

Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,

Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayaishizaki,Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

wagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,

Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,

Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.

and Hayaishizaki,Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)

Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,

Sugahara,Y. and Hayaishizaki,Y.

Computer-based methods for the mouse full-length cDNA

encyclopedia: real-time sequence clustering for construction of a

nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I.,

Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and

Hayaishizaki,Y.

Computational Analysis of Full-length Mouse cDNAs Compared with

Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for

FEATURES

1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:chimbesc@sc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)

Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.

PRIMERS

Sequencing: -21M13

LIBRARY

Vector : pKS145

R.Site 1 : SacI

R.Site 2 : SacI

Location/Qualifiers

1..722

/organism="Pan troglodytes"

/mol_type="genomic DNA"

/db_xref="taxon:9598"

/clone="PTB-069E13.F"

/sex="male"

/cell_type="lymphoblast"

/clone_lib="PTB Chimpanzee Male BAC Library"

ORIGIN

Alignment Scores:

Pred. No.:	582	Length:	722
Score:	43.00	Matches:	7
Percent Similarity:	78.57%	Conservative:	4
Best Local Similarity:	50.00%	Mismatches:	3
Query Match:	58.11%	Indels:	0
DB:	10	Gaps:	0

US-10-787-382-20 (1-15) x AG075393 (1-722)

Qy 2 AlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15

Db 385 GCAATTCAAGAACCTCAGAACAGAAAAATATCAGAAACATTG 426

Search completed: December 21, 2005, 18:53:14

Job time : 2639 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 21, 2005, 16:41:26 ; Search time 138 Seconds

(without alignments)
193.213 Million cell updates/sec

Title: US-10-787-382-20

Perfect score: 74

Sequence: 1 FAVENPMRLVAETL 15

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO spool/US10787382/runat 16122005 132436 8038/app_query.fasta_1.199
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCUI=0
-LOOPEXT=0 -UNITFS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=500 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=50
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10787382 @CGN 1 1 193 @runat 16122005 132436 8038 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/1 COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5 COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/H COMB.seq.*
6: /cgn2_6/ptodata/1/ina/PTCUS COMB.seq.*
7: /cgn2_6/ptodata/1/ina/pp COMB.seq.*
8: /cgn2_6/ptodata/1/ina/RE COMB.seq.*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	74	100.0	345	3	US-09-322-409-85
2	74	100.0	345	3	US-09-322-409-87
3	74	100.0	345	3	US-09-451-527-85
4	74	100.0	345	3	US-09-451-527-87
5	74	100.0	402	3	US-09-322-409-83
6	74	100.0	402	3	US-09-322-409-84
7	74	100.0	402	3	US-09-322-409-81
8	74	100.0	402	3	US-09-451-527-84
9	74	100.0	405	3	US-09-371-615A-1

Sequence 80, Appl	US-09-322-409-80	610	3	Sequence 80, Appl
Sequence 82, Appl	US-09-322-409-82	610	3	Sequence 82, Appl
Sequence 80, Appl	US-09-451-527-80	610	3	Sequence 80, Appl
Sequence 82, Appl	US-09-451-527-82	610	3	Sequence 82, Appl
Sequence 8759, Ap	US-09-248-796A-8759	210	3	Sequence 8759, Ap
Sequence 3, Appli	US-09-312-762A-3	14707	3	Sequence 3, Appli
Sequence 15193, A	US-09-949-016-15393	194790	3	Sequence 15193, A
Sequence 3240, Ap	US-09-949-016-3240	2085	3	Sequence 3240, Ap
Sequence 14982, A	US-09-949-016-14982	127771	3	Sequence 14982, A
Sequence 14033, A	US-09-949-016-14033	54.1	3	Sequence 14033, A
Sequence 12777, A	US-09-949-016-12777	828152	3	Sequence 12777, A
Sequence 23, Appl	US-08-680-326-23	3027	2	Sequence 23, Appl
Sequence 138, App	US-08-434-001-138	52.7	71	Sequence 138, App
Sequence 138, App	US-08-433-585-138	52.7	71	Sequence 138, App
Sequence 138, App	US-08-434-425-138	52.7	71	Sequence 138, App
Sequence 138, App	US-08-437-667-138	52.7	71	Sequence 138, App
Sequence 138, App	US-08-906-955-138	52.7	71	Sequence 138, App
Sequence 138, App	US-08-945-909-138	52.7	71	Sequence 138, App
Sequence 138, App	US-08-396-002A-138	52.7	71	Sequence 138, App
Sequence 138, App	US-10-077-319-138	52.7	71	Sequence 138, App
Sequence 138, App	PCT-US96-06060-138	52.7	71	Sequence 138, App
Sequence 8641, Ap	US-09-270-767-8641	523	3	Sequence 8641, Ap
Sequence 23923, A	US-09-270-767-23923	523	3	Sequence 23923, A
Sequence 45191, A	US-09-949-016-45191	601	3	Sequence 45191, A
Sequence 86, Appl	US-08-222-575-86	806	3	Sequence 86, Appl
Sequence 72, Appl	US-09-389-681-72	806	3	Sequence 72, Appl
Sequence 86, Appl	US-09-389-681-86	806	3	Sequence 86, Appl
Sequence 72, Appl	US-09-620-405B-72	806	3	Sequence 72, Appl
Sequence 86, Appl	US-09-620-405B-86	806	3	Sequence 86, Appl
Sequence 72, Appl	US-09-339-338-72	806	3	Sequence 72, Appl
Sequence 86, Appl	US-09-339-338-86	806	3	Sequence 86, Appl
Sequence 72, Appl	US-09-433-826B-72	806	3	Sequence 72, Appl
Sequence 86, Appl	US-09-433-826B-86	806	3	Sequence 86, Appl
Sequence 72, Appl	US-09-604-287A-72	806	3	Sequence 72, Appl
Sequence 86, Appl	US-09-604-287A-86	806	3	Sequence 86, Appl
Sequence 72, Appl	US-09-285-480-72	806	3	Sequence 72, Appl
Sequence 86, Appl	US-09-285-480-86	806	3	Sequence 86, Appl
Sequence 72, Appl	US-09-834-759-72	806	3	Sequence 72, Appl
Sequence 86, Appl	US-09-834-759-86	806	3	Sequence 86, Appl
Sequence 72, Appl	US-09-590-751A-72	806	3	Sequence 72, Appl
Sequence 86, Appl	US-09-590-751A-86	806	3	Sequence 86, Appl
Sequence 72, Appl	US-09-551-621-72	806	3	Sequence 72, Appl
Sequence 86, Appl	US-09-551-621-86	806	3	Sequence 86, Appl
Sequence 72, Appl	US-09-551-621A-72	806	3	Sequence 72, Appl
Sequence 86, Appl	US-09-551-621A-86	806	3	Sequence 86, Appl
Sequence 72, Appl	US-10-076-622-72	806	3	Sequence 72, Appl
Sequence 86, Appl	US-10-076-622-86	806	3	Sequence 86, Appl
Sequence 987, App	US-09-134-000C-987	1461	3	Sequence 987, App
Sequence 3069, Ap	US-09-543-681A-3069	1680	3	Sequence 3069, Ap
Sequence 16426, A	US-09-820-002-3	4897	3	Sequence 16426, A
Sequence 13053, A	US-09-949-016-13053	21784	3	Sequence 13053, A
Sequence 17010, A	US-09-949-016-17010	30137	3	Sequence 17010, A
Sequence 3, Appli	US-09-801-861-3	53332	3	Sequence 3, Appli
Sequence 3, Appli	US-10-224-062-3	53332	3	Sequence 3, Appli
Sequence 3, Appli	US-10-786-065-3	53332	3	Sequence 3, Appli
Sequence 13034, A	US-09-949-016-13034	90900	3	Sequence 13034, A
Sequence 16908, A	US-09-513-999C-16908	200	3	Sequence 16908, A
Sequence 153, App	US-09-614-221A-153	345	3	Sequence 153, App
Sequence 22319, A	US-09-949-016-22319	601	3	Sequence 22319, A
Sequence 22240, A	US-09-949-016-22240	601	3	Sequence 22240, A
Sequence 51346, A	US-09-949-016-51346	601	3	Sequence 51346, A
Sequence 51347, A	US-09-949-016-51347	601	3	Sequence 51347, A
Sequence 175044, A	US-09-949-016-175044	601	3	Sequence 175044, A
Sequence 196318, A	US-09-949-016-196318	601	3	Sequence 196318, A
Sequence 6645, Ap	US-09-902-540-6645	1044	3	Sequence 6645, Ap
Sequence 12117, A	US-09-252-991A-12117	1456	3	Sequence 12117, A
Sequence 2830, Ap	US-09-252-991A-2830	1425	3	Sequence 2830, Ap
Sequence 9746, Ap	US-09-252-991A-9746	1521	3	Sequence 9746, Ap
Sequence 970, App	US-09-134-000C-970	1650	3	Sequence 970, App
Sequence 12258, A	US-09-252-991A-12258	1656	3	Sequence 12258, A
Sequence 3, Appli	US-09-591-095-3	1992	3	Sequence 3, Appli
Sequence 9494, Ap	US-09-252-991A-9494	2511	3	Sequence 9494, Ap

83	38	51.4	2583	3	US-09-252-991A-9541	Sequence 9541, Ap	C 156	37	50.0	12366	3	US-09-949-016-13082	Sequence 13082, A
84	38	51.4	3131	3	US-09-902-540-513	Sequence 513, App	C 157	37	50.0	12619	3	US-09-616-289-49	Sequence 49, Appl
85	38	51.4	3618	3	US-09-252-991A-12305	Sequence 12305, A	C 158	37	50.0	12619	3	US-09-976-750-49	Sequence 14, Appl
C 86	38	51.4	8509	3	US-09-826-205-1	Sequence 1, Appl	C 159	37	50.0	12671	3	US-08-976-259-14	Sequence 14, Appl
C 87	38	51.4	10945	3	US-09-949-016-14097	Sequence 14097, A	C 160	37	50.0	22671	3	US-09-956-004-14	Sequence 14, Appl
C 88	38	51.4	87323	3	US-09-949-016-13828	Sequence 13828, A	C 161	37	50.0	24954	3	US-09-949-016-13475	Sequence 13475, A
C 89	38	51.4	121427	3	US-09-949-016-11950	Sequence 11950, A	C 162	37	50.0	42917	3	US-09-949-016-13795	Sequence 13795, A
C 90	38	51.4	121433	3	US-09-949-016-13230	Sequence 13230, A	C 163	37	50.0	44377	3	US-09-949-016-11840	Sequence 11840, A
C 91	38	51.4	159863	3	US-09-949-016-14858	Sequence 14858, A	C 164	37	50.0	44378	3	US-09-949-016-13569	Sequence 15669, A
C 92	38	51.4	171130	3	US-09-949-016-14861	Sequence 14861, A	C 165	37	50.0	49416	3	US-09-949-016-15234	Sequence 15234, A
C 93	38	51.4	174170	3	US-09-949-016-14810	Sequence 14810, A	C 166	37	50.0	70313	3	US-09-949-002-714	Sequence 714, App
C 94	38	51.4	174318	3	US-09-949-016-14811	Sequence 14811, A	C 167	37	50.0	71278	3	US-09-949-016-11851	Sequence 11851, A
C 95	38	51.4	174318	3	US-09-949-016-14812	Sequence 14812, A	C 168	37	50.0	71278	3	US-09-949-016-17563	Sequence 17563, A
C 96	38	51.4	174318	3	US-09-949-016-14813	Sequence 14813, A	C 169	37	50.0	72347	3	US-09-949-002-633	Sequence 633, App
C 97	38	51.4	451924	3	US-09-949-016-12896	Sequence 12896, A	C 170	37	50.0	72602	3	US-09-949-016-14385	Sequence 14385, A
C 98	38	51.4	451924	3	US-09-949-016-12896	Sequence 12896, A	C 171	37	50.0	84252	3	US-09-949-016-17315	Sequence 17315, A
C 99	38	51.4	451925	3	US-09-949-016-17305	Sequence 17305, A	C 172	37	50.0	86116	3	US-09-949-016-14766	Sequence 14766, A
C 100	38	51.4	462589	3	US-09-949-016-12900	Sequence 12900, A	C 173	37	50.0	168174	3	US-08-793-701-26	Sequence 63, Appl
C 101	38	51.4	476044	3	US-09-949-016-12412	Sequence 12412, A	C 174	37	50.0	168273	3	US-10-071-411A-63	Sequence 2, Appl
C 102	38	51.4	1664976	3	US-08-916-421B-1	Sequence 1, Appl	C 175	37	50.0	256171	3	US-09-949-016-12822	Sequence 12822, A
C 103	38	51.4	1664976	3	US-09-322-409-137	Sequence 137, App	C 176	37	50.0	256171	3	US-09-949-016-15524	Sequence 15524, A
C 104	37	50.0	36	3	US-09-451-527-137	Sequence 137, App	C 177	37	50.0	455726	3	US-09-949-016-14157	Sequence 14157, A
C 105	37	50.0	310	3	US-09-121-211-17	Sequence 17, Appl	C 178	37	50.0	481115	3	US-09-949-016-11940	Sequence 11940, A
C 106	37	50.0	601	3	US-09-949-016-19668	Sequence 19668, A	C 179	36	48.6	96	3	US-08-793-701-26	Sequence 26, Appl
C 107	37	50.0	601	3	US-09-949-016-150330	Sequence 150330, A	C 180	36	48.6	96	3	US-09-579-264-26	Sequence 26, Appl
C 108	37	50.0	601	3	US-09-949-016-150330	Sequence 150330, A	C 181	36	48.6	234	3	US-09-543-681A-3632	Sequence 3632, Ap
C 109	37	50.0	601	3	US-09-949-016-196769	Sequence 196769, A	C 182	36	48.6	249	2	US-08-687-080-100	Sequence 100, Appl
C 110	37	50.0	651	3	US-09-533-559-1764	Sequence 1764, Ap	C 183	36	48.6	295	3	US-08-905-223-74	Sequence 74, Appl
C 111	37	50.0	651	3	US-09-270-767-17025	Sequence 17025, A	C 184	36	48.6	318	3	US-09-248-796A-9413	Sequence 9413, Ap
C 112	37	50.0	726	3	US-10-187-790A-15	Sequence 15, Appl	C 185	36	48.6	358	3	US-09-702-705-1240	Sequence 1240, Ap
C 113	37	50.0	831	3	US-09-902-540-8561	Sequence 8561, Ap	C 186	36	48.6	358	3	US-09-736-457-1240	Sequence 1240, Ap
C 114	37	50.0	833	6	PCT-US91-08177-16	Sequence 16, Appl	C 187	36	48.6	358	3	US-09-614-124B-1240	Sequence 1240, Ap
C 115	37	50.0	835	3	US-09-533-559-1764	Sequence 1764, Ap	C 188	36	48.6	358	3	US-09-671-325-1240	Sequence 1240, Ap
C 116	37	50.0	1000	3	US-09-641-638-460	Sequence 460, App	C 189	36	48.6	358	3	US-09-658-824-1240	Sequence 1240, Ap
C 117	37	50.0	1000	3	US-10-170-097-460	Sequence 460, App	C 190	36	48.6	358	3	US-10-017-754-1240	Sequence 1240, Ap
C 118	37	50.0	1149	3	US-09-533-559-703	Sequence 703, App	C 191	36	48.6	358	3	US-09-651-563-1240	Sequence 1240, Ap
C 119	37	50.0	1433	2	US-07-968-971A-11	Sequence 11, Appl	C 192	36	48.6	372	3	US-09-583-110-2257	Sequence 2257, Ap
C 120	37	50.0	1433	2	US-08-383-756-5	Sequence 5, Appl	C 193	36	48.6	372	3	US-09-513-999C-29856	Sequence 29856, A
C 121	37	50.0	1433	2	US-08-424-408-2	Sequence 2, Appl	C 194	36	48.6	384	3	US-09-543-681A-3672	Sequence 3672, Ap
C 122	37	50.0	1433	2	US-08-464-523B-8	Sequence 8, Appl	C 195	36	48.6	406	3	US-09-270-767-6602	Sequence 6602, Ap
C 123	37	50.0	1433	2	US-08-464-523B-8	Sequence 5, Appl	C 196	36	48.6	406	3	US-09-270-767-21884	Sequence 21884, A
C 124	37	50.0	1521	3	US-09-252-991A-1594	Sequence 1594, Ap	C 197	36	48.6	409	3	US-09-513-999C-3638	Sequence 3638, Ap
C 125	37	50.0	1524	3	US-10-012-762-14	Sequence 14, Appl	C 198	36	48.6	552	2	US-08-602-010A-3	Sequence 3, Appl
C 126	37	50.0	1524	3	US-09-704-036B-14	Sequence 14, Appl	C 199	36	48.6	552	2	US-08-680-726A-3	Sequence 3, Appl
C 127	37	50.0	1524	3	US-10-651-183-14	Sequence 14, Appl	C 200	36	48.6	552	3	US-09-094-409-3	Sequence 3, Appl
C 128	37	50.0	1534	3	US-08-629-643A-4	Sequence 4, Appl	C 201	36	48.6	591	3	US-09-328-352-813	Sequence 813, App
C 129	37	50.0	1534	3	US-09-155-884-4	Sequence 4, Appl	C 202	36	48.6	591	3	US-09-107-433-1405	Sequence 1405, Ap
C 130	37	50.0	1674	3	US-09-252-991A-1645	Sequence 1645, Ap	C 203	36	48.6	601	3	US-09-949-016-50344	Sequence 50344, A
C 131	37	50.0	1761	3	US-09-519-232-7	Sequence 7, Appl	C 204	36	48.6	601	3	US-09-949-016-130270	Sequence 130270, A
C 132	37	50.0	1785	3	US-09-252-991A-1528	Sequence 1528, Ap	C 205	36	48.6	604	3	US-09-735-271-668	Sequence 668, App
C 133	37	50.0	2139	3	US-09-949-002-142	Sequence 142, App	C 206	36	48.6	606	3	US-09-711-164-138	Sequence 138, App
C 134	37	50.0	2583	3	US-10-104-047-1010	Sequence 1010, Ap	C 207	36	48.6	606	3	US-09-492-709A-232	Sequence 232, App
C 135	37	50.0	2677	3	US-09-270-767-28385	Sequence 28385, A	C 208	36	48.6	629	3	US-09-489-039A-3800	Sequence 3800, Ap
C 136	37	50.0	3106	3	US-09-323-872A-21	Sequence 21, Appl	C 209	36	48.6	629	3	US-10-178-213-46	Sequence 46, Appl
C 137	37	50.0	3106	3	US-09-072-433-13	Sequence 13, Appl	C 210	36	48.6	648	3	US-09-543-681A-3031	Sequence 3031, Ap
C 138	37	50.0	3360	3	US-09-121-211-1	Sequence 1, Appl	C 211	36	48.6	657	3	US-09-248-796A-126	Sequence 126, App
C 139	37	50.0	3420	2	US-08-117-491-25	Sequence 25, Appl	C 212	36	48.6	700	3	US-09-735-271-470	Sequence 470, App
C 140	37	50.0	3420	2	US-08-271-364A-6	Sequence 6, Appl	C 213	36	48.6	700	3	US-09-735-271-471	Sequence 471, App
C 141	37	50.0	3420	2	US-08-222-715B-25	Sequence 25, Appl	C 214	36	48.6	1047	3	US-09-221-017B-59	Sequence 59, Appl
C 142	37	50.0	4173	3	US-09-949-002-61	Sequence 61, Appl	C 215	36	48.6	1239	3	US-09-894-844-68	Sequence 68, Appl
C 143	37	50.0	4214	3	US-09-221-017B-293	Sequence 293, App	C 216	36	48.6	1266	3	US-09-533-559-5540	Sequence 5540, Ap
C 144	37	50.0	4320	3	US-09-134-000C-2728	Sequence 2728, Ap	C 217	36	48.6	1266	3	US-09-489-039A-6230	Sequence 6230, Ap
C 145	37	50.0	4707	2	US-08-004-139B-2	Sequence 2, Appl	C 218	36	48.6	1266	3	US-08-581-148C-19	Sequence 19, Appl
C 146	37	50.0	4707	2	US-08-811-492-2	Sequence 2, Appl	C 219	36	48.6	1507	2	US-07-906-871-13	Sequence 13, Appl
C 147	37	50.0	4707	6	PCT-US96-10545A-2	Sequence 2, Appl	C 220	36	48.6	1507	2	US-10-104-047-921	Sequence 921, Appl
C 148	37	50.0	4928	3	US-09-323-872A-32	Sequence 32, Appl	C 221	36	48.6	1923	3	US-09-673-395A-46	Sequence 46, Appl
C 149	37	50.0	4928	3	US-09-072-433-17	Sequence 17, Appl	C 222	36	48.6	2095	3	US-10-104-047-1393	Sequence 1393, Ap
C 150	37	50.0	5706	3	US-09-270-767-12590	Sequence 12590, A	C 223	36	48.6	2146	3	US-09-620-312D-443	Sequence 443, App
C 151	37	50.0	6727	3	US-08-629-643A-5	Sequence 5, Appl	C 224	36	48.6	2146	3	US-09-949-016-4667	Sequence 1067, Ap
C 152	37	50.0	6727	3	US-09-280-799-1	Sequence 1, Appl	C 225	36	48.6	2451	3	US-09-489-039A-3449	Sequence 3449, Ap
C 153	37	50.0	6727	3	US-09-155-884-5	Sequence 5, Appl	C 226	36	48.6	2523	3	US-09-489-039A-3473	Sequence 3473, Ap
C 154	37	50.0	9325	3	US-09-902-540-903	Sequence 903, App	C 227	36	48.6	2529	3	US-09-248-796A-1502	Sequence 1502, Ap
C 155	37	50.0	9482	3	US-09-949-016-14605	Sequence 14605, A	C 228	36	48.6	2631	2	US-08-717-515-3	Sequence 3, Appl

C 229	36	48.6	3255	2	US-08-717-515-5	Sequence 5, Appli	302	35	47.3	579	3	US-09-252-991A-15036	Sequence 15036, A
C 230	36	48.6	3279	6	PCT-US93-03077-2	Sequence 2, Appli	303	35	47.3	585	3	US-09-107-433-1045	Sequence 1045, Ap
C 231	36	48.6	3282	3	US-09-949-016-5664	Sequence 5664, Ap	304	35	47.3	594	3	US-09-902-540-7441	Sequence 7441, Ap
C 232	36	48.6	3831	2	US-08-717-515-7	Sequence 7, Appli	C 305	35	47.3	601	3	US-09-949-016-26797	Sequence 26797, A
C 233	36	48.6	4307	3	US-09-949-002-94	Sequence 94, Appli	C 306	35	47.3	601	3	US-09-949-016-28298	Sequence 28298, A
C 234	36	48.6	4407	3	US-10-012-231A-316	Sequence 316, App	C 307	35	47.3	601	3	US-09-949-016-28299	Sequence 28299, A
C 235	36	48.6	4407	3	US-10-013-389A-316	Sequence 316, App	C 308	35	47.3	601	3	US-09-949-016-33343	Sequence 33343, A
C 236	36	48.6	4407	3	US-10-006-768A-316	Sequence 316, App	C 309	35	47.3	601	3	US-09-949-016-33344	Sequence 33344, A
C 237	36	48.6	4407	3	US-10-015-671A-316	Sequence 316, App	C 310	35	47.3	601	3	US-09-949-016-68235	Sequence 68235, A
C 238	36	48.6	4407	3	US-10-015-393A-316	Sequence 316, App	C 311	35	47.3	601	3	US-09-949-016-159686	Sequence 159686, A
C 239	36	48.6	4407	3	US-10-011-833A-316	Sequence 316, App	C 312	35	47.3	601	3	US-09-949-016-169149	Sequence 169149, A
C 240	36	48.6	4407	3	US-10-006-041A-316	Sequence 316, App	C 313	35	47.3	601	3	US-09-949-016-171968	Sequence 171968, A
C 241	36	48.6	4407	3	US-10-012-064A-316	Sequence 316, App	C 314	35	47.3	601	3	US-09-949-016-182563	Sequence 182563, A
C 242	36	48.6	4562	3	US-09-620-312D-418	Sequence 418, App	C 315	35	47.3	601	3	US-09-949-016-182564	Sequence 182564, A
C 243	36	48.6	5495	2	US-08-602-010A-1	Sequence 1, Appli	C 316	35	47.3	601	3	US-09-949-016-182683	Sequence 182683, A
C 244	36	48.6	5495	2	US-08-602-010A-2	Sequence 2, Appli	C 317	35	47.3	601	3	US-09-949-016-182684	Sequence 182684, A
C 245	36	48.6	5495	2	US-08-680-726A-1	Sequence 1, Appli	C 318	35	47.3	601	3	US-09-949-016-188683	Sequence 188683, A
C 246	36	48.6	5495	2	US-08-680-726A-2	Sequence 2, Appli	C 319	35	47.3	601	3	US-09-949-016-188683	Sequence 188683, A
C 247	36	48.6	5495	3	US-09-092-409-1	Sequence 1, Appli	C 320	35	47.3	601	3	US-09-949-016-188732	Sequence 188732, A
C 248	36	48.6	5495	3	US-09-092-409-2	Sequence 2, Appli	C 321	35	47.3	601	3	US-09-949-016-188781	Sequence 188781, A
C 249	36	48.6	6294	3	US-08-976-259-5	Sequence 5, Appli	C 322	35	47.3	601	3	US-09-949-016-192865	Sequence 192865, A
C 250	36	48.6	6294	3	US-09-956-004-5	Sequence 5, Appli	C 323	35	47.3	601	3	US-09-949-016-192911	Sequence 192911, A
C 251	36	48.6	6830	2	US-08-822-445-1	Sequence 1, Appli	C 324	35	47.3	601	3	US-09-949-016-192957	Sequence 192957, A
C 252	36	48.6	6830	2	US-09-396-540-1	Sequence 1, Appli	C 325	35	47.3	601	3	US-09-949-016-193003	Sequence 193003, A
C 253	36	48.6	6866	3	US-09-949-016-15807	Sequence 15807, A	C 326	35	47.3	601	3	US-09-949-016-193049	Sequence 193049, A
C 254	36	48.6	10592	2	US-08-680-726A-51	Sequence 51, Appli	C 327	35	47.3	601	3	US-09-949-016-193142	Sequence 193142, A
C 255	36	48.6	10592	2	US-08-680-726A-52	Sequence 52, Appli	C 328	35	47.3	601	3	US-09-949-016-193235	Sequence 193235, A
C 256	36	48.6	10592	3	US-09-092-409-51	Sequence 51, Appli	C 329	35	47.3	601	3	US-09-949-016-193328	Sequence 193328, A
C 257	36	48.6	10592	3	US-09-092-409-52	Sequence 52, Appli	C 330	35	47.3	606	3	US-09-543-681A-3568	Sequence 3568, Ap
C 258	36	48.6	11073	3	US-09-949-016-15809	Sequence 15809, A	C 331	35	47.3	615	3	US-09-134-000C-674	Sequence 674, App
C 259	36	48.6	11165	3	US-09-830-807-34	Sequence 34, Appli	C 332	35	47.3	615	3	US-10-012-819-195	Sequence 195, App
C 260	36	48.6	12225	2	US-08-822-445-11	Sequence 11, Appli	C 333	35	47.3	648	3	US-09-543-681A-152	Sequence 152, App
C 261	36	48.6	12225	3	US-09-396-540-11	Sequence 11, Appli	C 334	35	47.3	741	3	US-09-107-532A-154	Sequence 154, App
C 262	36	48.6	12570	3	US-09-949-002-850	Sequence 850, App	C 335	35	47.3	745	3	US-08-858-207A-83	Sequence 83, Appli
C 263	36	48.6	12616	2	US-08-822-445-9	Sequence 9, Appli	C 336	35	47.3	768	3	US-09-252-991A-14591	Sequence 14591, A
C 264	36	48.6	20116	3	US-09-396-540-9	Sequence 9, Appli	C 337	35	47.3	771	3	US-09-252-991A-13903	Sequence 13903, A
C 265	36	48.6	20407	3	US-09-949-016-16861	Sequence 16861, A	C 338	35	47.3	777	3	US-09-252-991A-13903	Sequence 13903, A
C 266	36	48.6	20407	3	US-09-949-002-666	Sequence 666, App	C 339	35	47.3	780	3	US-09-489-039A-1590	Sequence 1590, Ap
C 267	36	48.6	32868	3	US-09-949-016-17406	Sequence 17406, A	C 340	35	47.3	828	3	US-09-489-039A-5862	Sequence 5862, Ap
C 268	36	48.6	47375	3	US-09-949-016-15420	Sequence 15420, A	C 341	35	47.3	849	3	US-09-107-433-842	Sequence 842, Ap
C 269	36	48.6	62311	3	US-09-949-016-14582	Sequence 14582, A	C 342	35	47.3	849	3	US-09-252-991A-15171	Sequence 15171, A
C 270	36	48.6	92304	3	US-09-949-016-15943	Sequence 15943, A	C 343	35	47.3	978	3	US-09-489-039A-2498	Sequence 2498, Ap
C 271	36	48.6	117391	3	US-09-949-016-13945	Sequence 13945, A	C 344	35	47.3	1080	3	US-09-248-796A-3766	Sequence 3766, Ap
C 272	36	48.6	126237	3	US-09-949-016-16674	Sequence 16674, A	C 345	35	47.3	1185	3	US-08-705-771-10	Sequence 10, Appli
C 273	36	48.6	126237	3	US-09-949-016-16675	Sequence 16675, A	C 346	35	47.3	1185	3	US-09-417-540-10	Sequence 10, Appli
C 274	36	48.6	168104	3	US-09-949-016-12026	Sequence 12026, A	C 347	35	47.3	1200	3	US-09-602-787A-465	Sequence 465, App
C 275	36	48.6	168105	3	US-09-949-016-15554	Sequence 15554, A	C 348	35	47.3	1278	3	US-09-252-991A-14456	Sequence 14456, A
C 276	36	48.6	198942	3	US-09-949-016-13209	Sequence 13209, A	C 349	35	47.3	1314	3	US-09-248-796A-3402	Sequence 3402, Ap
C 277	36	48.6	232024	3	US-09-949-016-13477	Sequence 13477, A	C 350	35	47.3	1334	3	US-09-902-540-8599	Sequence 8599, Ap
C 278	36	48.6	786431	3	US-09-751-389-3	Sequence 3, Appli	C 351	35	47.3	1368	3	US-09-540-236-405	Sequence 405, App
C 279	36	48.6	8403765	3	US-09-103-840A-2	Sequence 2, Appli	C 352	35	47.3	1419	3	US-09-252-991A-13796	Sequence 13796, A
C 280	36	48.6	4411529	3	US-09-103-840A-1	Sequence 1, Appli	C 353	35	47.3	1583	3	US-09-620-312D-532	Sequence 532, App
C 281	35	47.3	113	3	US-09-513-998C-33867	Sequence 33867, A	C 354	35	47.3	1599	3	US-09-902-540-7144	Sequence 7144, Ap
C 282	35	47.3	228	3	US-09-107-433-766	Sequence 766, App	C 355	35	47.3	1611	2	US-08-551-211-4	Sequence 4, Appli
C 283	35	47.3	249	3	US-09-270-767-3988	Sequence 3988, Ap	C 356	35	47.3	1771	3	US-09-270-767-13467	Sequence 13467, A
C 284	35	47.3	249	3	US-09-270-767-13270	Sequence 13270, Ap	C 357	35	47.3	1771	3	US-09-495-797-40	Sequence 40, Appli
C 285	35	47.3	288	3	US-09-543-681A-3498	Sequence 3498, Ap	C 358	35	47.3	1773	3	US-09-533-559-4781	Sequence 4781, Ap
C 286	35	47.3	454	3	US-09-902-540-6951	Sequence 6951, Ap	C 359	35	47.3	1823	3	US-08-553-619B-1	Sequence 1, Appli
C 287	35	47.3	456	3	US-09-489-039A-4007	Sequence 4007, Ap	C 360	35	47.3	1860	2	US-08-553-619B-1	Sequence 1, Appli
C 288	35	47.3	471	3	US-09-668-262A-5	Sequence 5, Appli	C 361	35	47.3	1866	3	US-08-668-262A-15	Sequence 15, Appli
C 289	35	47.3	471	3	US-10-427-442-5	Sequence 5, Appli	C 362	35	47.3	1866	3	US-10-427-442-15	Sequence 15, Appli
C 290	35	47.3	493	3	US-09-702-705-750	Sequence 750, App	C 363	35	47.3	1908	3	US-09-252-991A-14442	Sequence 14442, A
C 291	35	47.3	493	3	US-09-736-457-750	Sequence 750, App	C 364	35	47.3	1995	3	US-09-252-991A-13546	Sequence 13546, A
C 292	35	47.3	493	3	US-09-614-1248-750	Sequence 750, App	C 365	35	47.3	2040	3	US-09-527-972-3	Sequence 3, Appli
C 293	35	47.3	493	3	US-09-671-325-750	Sequence 750, App	C 366	35	47.3	2040	3	US-08-581-148C-21	Sequence 21, Appli
C 294	35	47.3	493	3	US-09-589-184-750	Sequence 750, App	C 367	35	47.3	2121	3	US-10-104-047-771	Sequence 771, App
C 295	35	47.3	493	3	US-09-658-821-750	Sequence 750, App	C 368	35	47.3	2202	2	US-08-551-211-5	Sequence 5, Appli
C 296	35	47.3	493	3	US-10-017-754-750	Sequence 750, App	C 369	35	47.3	2367	3	US-09-583-110-1278	Sequence 1278, Ap
C 297	35	47.3	493	3	US-09-651-563-750	Sequence 750, App	C 370	35	47.3	2539	3	US-08-581-148C-20	Sequence 20, Appli
C 298	35	47.3	493	3	US-09-519-642-750	Sequence 750, App	C 371	35	47.3	2571	3	US-09-107-433-931	Sequence 931, App
C 299	35	47.3	522	3	US-09-248-796A-2604	Sequence 2604, Ap	C 372	35	47.3	2621	2	US-08-553-619B-8	Sequence 8, Appli
C 300	35	47.3	552	3	US-09-621-976-150	Sequence 150, App	C 373	35	47.3	3215	3	US-09-155-885A-299	Sequence 299, App
C 301	35	47.3	558	3	US-09-583-110-2493	Sequence 2493, Ap	C 374	35	47.3	3446	3	US-09-620-312D-653	Sequence 653, App


```
; NUMBER OF SEQ ID NOS: 154
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 85
; LENGTH: 345
; TYPE: DNA
; ORGANISM: Canis familiaris
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(345)
US-09-322-409-85

Alignment Scores:
Pred. No.: 9.69e-07 Length: 345
Score: 74.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-787-382-20 (1-15) x US-09-322-409-85 (1-345)
Qy 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
Db 1 TTTGCTGTAGAAAATCCCATGAACTAGACTGGTGGCAGACCTTG 45

RESULT 2
US-09-322-409-87/c
; Sequence 87, Application US/09322409
; Patent No. 6471957
; GENERAL INFORMATION:
; APPLICANT: Sim, Gek-Kee
; APPLICANT: Yang, Shumin
; APPLICANT: Dreitz, Matthew J.
; APPLICANT: Wonderling, Ramani S.
; TITLE OF INVENTION: CANINE AND FELINE IMMUNOREGULATORY PROTEINS, NUCLEIC
; FILE REFERENCE: IM-2-C1
; CURRENT APPLICATION NUMBER: US/09/322,409
; EARLIER FILING DATE: 1999-05-28
; EARLIER APPLICATION NUMBER: 60/087,306
; EARLIER FILING DATE: 1998-05-29
; NUMBER OF SEQ ID NOS: 154
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 87
; LENGTH: 345
; TYPE: DNA
; ORGANISM: Canis familiaris
US-09-322-409-87

Alignment Scores:
Pred. No.: 9.69e-07 Length: 345
Score: 74.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-787-382-20 (1-15) x US-09-322-409-87 (1-345)
Qy 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
Db 345 TTTGCTGTAGAAAATCCCATGAACTAGACTGGTGGCAGACCTTG 301

RESULT 3
US-09-451-527-85
; Sequence 85, Application US/09451527
; Patent No. 6482403
; GENERAL INFORMATION:
; APPLICANT: Sim, Gek-Kee
; APPLICANT: Yang, Shumin
; APPLICANT: Dreitz, Matthew J.
; APPLICANT: Wonderling, Ramani S.
; TITLE OF INVENTION: CANINE AND FELINE IMMUNOREGULATORY PROTEINS, NUCLEIC
```

```
; TITLE OF INVENTION: ACID MOLECULES, AND USES THEREOF
; FILE REFERENCE: IM-2-C2
; CURRENT APPLICATION NUMBER: US/09/451,527
; CURRENT FILING DATE: 1999-12-01
; EARLIER APPLICATION NUMBER: 09/322,409
; EARLIER FILING DATE: 1999-05-28
; EARLIER APPLICATION NUMBER: 60/087,306
; EARLIER FILING DATE: 1998-05-29
; NUMBER OF SEQ ID NOS: 174
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 85
; LENGTH: 345
; TYPE: DNA
; ORGANISM: Canis familiaris
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(345)
US-09-451-527-85

Alignment Scores:
Pred. No.: 9.69e-07 Length: 345
Score: 74.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-787-382-20 (1-15) x US-09-451-527-85 (1-345)
Qy 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
Db 1 TTTGCTGTAGAAAATCCCATGAACTAGACTGGTGGCAGACCTTG 45

RESULT 4
US-09-451-527-87/c
; Sequence 87, Application US/09451527
; Patent No. 6482403
; GENERAL INFORMATION:
; APPLICANT: Sim, Gek-Kee
; APPLICANT: Yang, Shumin
; APPLICANT: Dreitz, Matthew J.
; APPLICANT: Wonderling, Ramani S.
; TITLE OF INVENTION: CANINE AND FELINE IMMUNOREGULATORY PROTEINS, NUCLEIC
; FILE REFERENCE: IM-2-C2
; CURRENT APPLICATION NUMBER: US/09/451,527
; CURRENT FILING DATE: 1999-12-01
; EARLIER APPLICATION NUMBER: 09/322,409
; EARLIER FILING DATE: 1999-05-28
; EARLIER APPLICATION NUMBER: 60/087,306
; EARLIER FILING DATE: 1998-05-29
; NUMBER OF SEQ ID NOS: 174
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 87
; LENGTH: 345
; TYPE: DNA
; ORGANISM: Canis familiaris
US-09-451-527-87

Alignment Scores:
Pred. No.: 9.69e-07 Length: 345
Score: 74.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-787-382-20 (1-15) x US-09-451-527-87 (1-345)
Qy 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
Db 345 TTTGCTGTAGAAAATCCCATGAACTAGACTGGTGGCAGACCTTG 301
```

```
RESULT 5
US-09-322-409-83
; Sequence 83, Application US/09322409
; Patent No. 6471957
; GENERAL INFORMATION:
; APPLICANT: Sim, Gek-Kee
; APPLICANT: Yang, Shumin
; APPLICANT: Dreitz, Matthew J.
; APPLICANT: Wonderling, Ramani S.
; TITLE OF INVENTION: CANINE AND FELINE IMMUNOREGULATORY PROTEINS, NUCLEIC
; FILE REFERENCE: IM-2-C1
; CURRENT APPLICATION NUMBER: US/09/322,409
; CURRENT FILING DATE: 1999-05-28
; EARLIER APPLICATION NUMBER: 60/087,306
; EARLIER FILING DATE: 1998-05-29
; NUMBER OF SEQ ID NOS: 154
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 83
; LENGTH: 402
; TYPE: DNA
; ORGANISM: Canis familiaris
US-09-322-409-83

Alignment Scores:
Pred. No.: 1.18e-06 Length: 402
Score: 74.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-787-382-20 (1-15) x US-09-322-409-83 (1-402)
QY 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
DB 58 TTTCGTGTAGAAAATCCCATGAATAGACTGGTGGCAGACCTTG 102

RESULT 6
US-09-322-409-84/c
; Sequence 84, Application US/09322409
; Patent No. 6471957
; GENERAL INFORMATION:
; APPLICANT: Sim, Gek-Kee
; APPLICANT: Yang, Shumin
; APPLICANT: Dreitz, Matthew J.
; APPLICANT: Wonderling, Ramani S.
; TITLE OF INVENTION: CANINE AND FELINE IMMUNOREGULATORY PROTEINS, NUCLEIC
; FILE REFERENCE: IM-2-C1
; CURRENT APPLICATION NUMBER: US/09/322,409
; CURRENT FILING DATE: 1999-05-28
; EARLIER APPLICATION NUMBER: 60/087,306
; EARLIER FILING DATE: 1998-05-29
; NUMBER OF SEQ ID NOS: 154
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 84
; LENGTH: 402
; TYPE: DNA
; ORGANISM: Canis familiaris
US-09-322-409-84

Alignment Scores:
Pred. No.: 1.18e-06 Length: 402
Score: 74.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-787-382-20 (1-15) x US-09-322-409-84 (1-402)
QY 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
DB 58 TTTCGTGTAGAAAATCCCATGAATAGACTGGTGGCAGACCTTG 102
```

```
Db 345 TTTCGTGTAGAAAATCCCATGAATAGACTGGTGGCAGACCTTG 301

RESULT 7
US-09-451-527-83
; Sequence 83, Application US/09451527
; Patent No. 6482403
; GENERAL INFORMATION:
; APPLICANT: Sim, Gek-Kee
; APPLICANT: Yang, Shumin
; APPLICANT: Dreitz, Matthew J.
; APPLICANT: Wonderling, Ramani S.
; TITLE OF INVENTION: CANINE AND FELINE IMMUNOREGULATORY PROTEINS, NUCLEIC
; FILE REFERENCE: IM-2-C2
; CURRENT APPLICATION NUMBER: US/09/451,527
; CURRENT FILING DATE: 1999-12-01
; EARLIER APPLICATION NUMBER: 09/322,409
; EARLIER FILING DATE: 1999-05-28
; EARLIER APPLICATION NUMBER: 60/087,306
; EARLIER FILING DATE: 1998-05-29
; NUMBER OF SEQ ID NOS: 174
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 83
; LENGTH: 402
; TYPE: DNA
; ORGANISM: Canis familiaris
US-09-451-527-83

Alignment Scores:
Pred. No.: 1.18e-06 Length: 402
Score: 74.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-787-382-20 (1-15) x US-09-451-527-83 (1-402)
QY 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
DB 58 TTTCGTGTAGAAAATCCCATGAATAGACTGGTGGCAGACCTTG 102

RESULT 8
US-09-451-527-84/c
; Sequence 84, Application US/09451527
; Patent No. 6482403
; GENERAL INFORMATION:
; APPLICANT: Sim, Gek-Kee
; APPLICANT: Yang, Shumin
; APPLICANT: Dreitz, Matthew J.
; APPLICANT: Wonderling, Ramani S.
; TITLE OF INVENTION: CANINE AND FELINE IMMUNOREGULATORY PROTEINS, NUCLEIC
; FILE REFERENCE: IM-2-C2
; CURRENT APPLICATION NUMBER: US/09/451,527
; CURRENT FILING DATE: 1999-12-01
; EARLIER APPLICATION NUMBER: 09/322,409
; EARLIER FILING DATE: 1999-05-28
; EARLIER APPLICATION NUMBER: 60/087,306
; EARLIER FILING DATE: 1998-05-29
; NUMBER OF SEQ ID NOS: 174
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 84
; LENGTH: 402
; TYPE: DNA
; ORGANISM: Canis familiaris
US-09-451-527-84

Alignment Scores:
Pred. No.: 1.18e-06 Length: 402
Score: 74.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-787-382-20 (1-15) x US-09-451-527-84 (1-402)
QY 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
DB 58 TTTCGTGTAGAAAATCCCATGAATAGACTGGTGGCAGACCTTG 102
```

```
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-787-382-20 (1-15) x US-09-451-527-84 (1-402)

Qy 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
Db 345 TTTGCTGTAGAAAATCCCATGAATAGACTGGTGGCAGAGACCTTG 301

RESULT 9
US-09-371-615A-1
; Sequence 1, Application US/09371615A
; Patent No. 6537781
; GENERAL INFORMATION:
; APPLICANT: IDEXX LABORATORIES
; TITLE OF INVENTION: METHODS AND COMPOSITIONS CONCERNING
; TITLE OF INVENTION: CANINE INTERLEUKIN 5
; FILE REFERENCE: 036040017000500
; CURRENT APPLICATION NUMBER: US/09/371,615A
; CURRENT FILING DATE: 1999-08-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PstSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 405
; TYPE: DNA
; ORGANISM: Canis familiaris
US-09-371-615A-1

Alignment Scores:
Pred. No.: 1.19e-06 Length: 405
Score: 74.00 Matches: 15
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-787-382-20 (1-15) x US-09-371-615A-1 (1-405)

Qy 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
Db 58 TTTGCTGTAGAAAATCCCATGAATAGACTGGTGGCAGAGACCTTG 102

RESULT 10
US-09-322-409-80
; Sequence 80, Application US/09322409
; Patent No. 6471957
; GENERAL INFORMATION:
; APPLICANT: Sim, Gek-kee
; APPLICANT: Yang, Shumin
; APPLICANT: Dreitz, Matthew J.
; APPLICANT: Wonderling, Ramani S.
; TITLE OF INVENTION: CANINE AND FELINE IMMUNOREGULATORY PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES, AND USES THEREOF
; FILE REFERENCE: IM-2-C1
; CURRENT APPLICATION NUMBER: US/09/322,409
; CURRENT FILING DATE: 1999-05-28
; EARLIER APPLICATION NUMBER: 60/087,306
; EARLIER FILING DATE: 1998-05-29
; NUMBER OF SEQ ID NOS: 154
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 80
; LENGTH: 610
; TYPE: DNA
; ORGANISM: Canis familiaris
; NAME/KEY: CDS
; LOCATION: (29) ..(430)
US-09-322-409-80

Alignment Scores:
Pred. No.: 2.01e-06 Length: 610
Score: 74.00 Matches: 15
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0
```

```
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-787-382-20 (1-15) x US-09-322-409-80 (1-610)

Qy 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
Db 86 TTTGCTGTAGAAAATCCCATGAATAGACTGGTGGCAGAGACCTTG 130

RESULT 11
US-09-322-409-82/c
; Sequence 82, Application US/09322409
; Patent No. 6471957
; GENERAL INFORMATION:
; APPLICANT: Sim, Gek-kee
; APPLICANT: Yang, Shumin
; APPLICANT: Dreitz, Matthew J.
; APPLICANT: Wonderling, Ramani S.
; TITLE OF INVENTION: CANINE AND FELINE IMMUNOREGULATORY PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES, AND USES THEREOF
; FILE REFERENCE: IM-2-C1
; CURRENT APPLICATION NUMBER: US/09/322,409
; CURRENT FILING DATE: 1999-05-28
; EARLIER APPLICATION NUMBER: 60/087,306
; EARLIER FILING DATE: 1998-05-29
; NUMBER OF SEQ ID NOS: 154
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 82
; LENGTH: 610
; TYPE: DNA
; ORGANISM: Canis familiaris
US-09-322-409-82

Alignment Scores:
Pred. No.: 2.01e-06 Length: 610
Score: 74.00 Matches: 15
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-787-382-20 (1-15) x US-09-322-409-82 (1-610)

Qy 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
Db 525 TTTGCTGTAGAAAATCCCATGAATAGACTGGTGGCAGAGACCTTG 481

RESULT 12
US-09-451-527-80
; Sequence 80, Application US/09451527
; Patent No. 6482403
; GENERAL INFORMATION:
; APPLICANT: Sim, Gek-kee
; APPLICANT: Yang, Shumin
; APPLICANT: Dreitz, Matthew J.
; APPLICANT: Wonderling, Ramani S.
; TITLE OF INVENTION: CANINE AND FELINE IMMUNOREGULATORY PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES, AND USES THEREOF
; FILE REFERENCE: IM-2-C2
; CURRENT APPLICATION NUMBER: US/09/451,527
; CURRENT FILING DATE: 1999-12-01
; EARLIER APPLICATION NUMBER: 09/322,409
; EARLIER FILING DATE: 1999-05-28
; EARLIER APPLICATION NUMBER: 60/087,306
; EARLIER FILING DATE: 1998-05-29
; NUMBER OF SEQ ID NOS: 174
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 80
; LENGTH: 610
; TYPE: DNA
; ORGANISM: Canis familiaris
```

```
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (29) ..(430)
US-09-451-527-80

Alignment Scores:
Pred. No.: 2,01e-06 Length: 610
Score: 74.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-787-382-20 (1-15) x US-09-451-527-80 (1-610)
Qy 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
Db 86 TTTCGTGTAGAAAATCCCATGAATAGACTGGTGGCAGACACCTTG 130

RESULT 13
US-09-451-527-82/c
/ Sequence 82, Application US/09451527
/ Patent No. 6482403
/ GENERAL INFORMATION:
/ APPLICANT: Sim, Gek-Kee
/ APPLICANT: Yang, Shumin
/ APPLICANT: Dreitz, Matthew J.
/ TITLE OF INVENTION: CANINE AND FELINE IMMUNOREGULATORY PROTEINS, NUCLEIC
/ TITLE OF INVENTION: ACID MOLECULES, AND USES THEREOF
/ FILE REFERENCE: IM-2-C2
/ CURRENT APPLICATION NUMBER: US/09/451,527
/ CURRENT FILING DATE: 1999-12-01
/ EARLIER APPLICATION NUMBER: 09/322,409
/ EARLIER FILING DATE: 1999-05-28
/ EARLIER APPLICATION NUMBER: 60/087,306
/ EARLIER FILING DATE: 1998-05-29
/ NUMBER OF SEQ ID NOS: 174
/ SOFTWARE: Patentin Ver. 2.0
/ SEQ ID NO 82
/ LENGTH: 610
/ TYPE: DNA
/ ORGANISM: Canis familiaris
US-09-451-527-82

Alignment Scores:
Pred. No.: 2,01e-06 Length: 610
Score: 74.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-787-382-20 (1-15) x US-09-451-527-82 (1-610)
Qy 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
Db 525 TTTCGTGTAGAAAATCCCATGAATAGACTGGTGGCAGACACCTTG 481

RESULT 14
US-09-248-796A-8759
/ Sequence 8759, Application US/09248796A
/ Patent No. 6747137
/ GENERAL INFORMATION:
/ APPLICANT: Keith Weinstein et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
/ TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: 107196.132
/ CURRENT APPLICATION NUMBER: US/09/248,796A
/ CURRENT FILING DATE: 1999-02-12
/ PRIOR APPLICATION NUMBER: US 60/074,725
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR APPLICATION NUMBER: US 60/096,409

/ PRIOR FILING DATE: 1998-08-13
/ NUMBER OF SEQ ID NOS: 28208
/ LENGTH: 210
/ TYPE: DNA
/ ORGANISM: Candida albicans
US-09-248-796A-8759

Alignment Scores:
Pred. No.: 2 Length: 210
Score: 43.00 Matches: 7
Percent Similarity: 90.91% Conservative: 3
Best Local Similarity: 63.64% Mismatches: 1
Query Match: 58.11% Indels: 0
DB: 3 Gaps: 0

US-10-787-382-20 (1-15) x US-09-248-796A-8759 (1-210)
Qy 3 ValGluAsnProMetAsnArgLeuValAlaGlu 13
Db 73 ATTGAATATCCCTTAACAGGTTGACAGCTAAA 105

RESULT 15
US-09-312-762A-3/c
/ Sequence 3, Application US/09312762A
/ Patent No. 6552177
/ GENERAL INFORMATION:
/ APPLICANT: MIA HOROWITZ ET AL.
/ TITLE OF INVENTION: EH DOMAIN CONTAINING GENES AND PROTEINS
/ NUMBER OF SEQUENCES: 27
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Mark M. Friedman c/o Anthony Castorina
/ STREET: 2001 Jefferson Davis Highway, Suite 207
/ CITY: Arlington
/ STATE: Virginia
/ COUNTRY: United States of America
/ ZIP: 22202
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk
/ COMPUTER: Twinhead* Slimnote-890TX
/ OPERATING SYSTEM: MS DOS version 6.2,
/ OPERATING SYSTEM: Windows version 3.11
/ SOFTWARE: Word for Windows version 2.0 converted to
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/312,762A
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 09/026,898
/ FILING DATE: 20 FEB 1998
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Friedman, Mark M.
/ REGISTRATION NUMBER: 33,883
/ REFERENCE/DOCKET NUMBER: 916/10
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 972-3-5625553
/ TELEFAX: 972-3-5625554
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 14707
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
US-09-312-762A-3

Alignment Scores:
Pred. No.: 753 Length: 14707
Score: 42.00 Matches: 8
Percent Similarity: 69.23% Conservative: 1
Best Local Similarity: 61.54% Mismatches: 4
Query Match: 56.76% Indels: 0
```

DB: 3 0 Gaps: 0
US-10-787-382-20 (1-15) x US-09-312-762A-3 (1-14707)
Qy 2 AlaValGluAsnProMetAsnArgLeuValAlaGluThr 14
Db 12589 GCTGCAGACATCTTCACGAGATTAAACAATGGAAACC 12551

RESULT 16

US-09-949-016-15393/c
; Sequence 15393, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15393
; LENGTH: 194790
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15393

Alignment Scores:

Pred. No.: 2,06e+04 Length: 194790
Score: 42.00 Matches: 8
Percent Similarity: 81.82% Conservative: 1
Best Local Similarity: 72.73% Mismatches: 2
Query Match: 56.76% Indels: 0
DB: 3 Gaps: 0

US-10-787-382-20 (1-15) x US-09-949-016-15393 (1-194790)

Qy 1 PheAlaValGluAsnProMetAsnArgLeuVal 11
Db 29935 TTTCCTGAGATCCAGTTGGAAGATTGGTA 29903

RESULT 17

US-09-949-016-3240/c
; Sequence 3240, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3240
; LENGTH: 2085
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-3240

Alignment Scores:

Pred. No.: 164 Length: 2085
Score: 40.00 Matches: 8
Percent Similarity: 69.23% Conservative: 1
Best Local Similarity: 61.54% Mismatches: 4
Query Match: 54.05% Indels: 0
DB: 3 Gaps: 0

US-10-787-382-20 (1-15) x US-09-949-016-3240 (1-2085)

Qy 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGlu 13
Db 2020 TTTCGGCTAGAAAAGCCCTTTTCAGAGGTTGGTGAGGAA 1982

RESULT 18

US-09-949-016-14982/c
; Sequence 14982, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14982
; LENGTH: 127771
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14982

Alignment Scores:

Pred. No.: 3.17e+04 Length: 127771
Score: 40.00 Matches: 8
Percent Similarity: 69.23% Conservative: 1
Best Local Similarity: 61.54% Mismatches: 4
Query Match: 54.05% Indels: 0
DB: 3 Gaps: 0

US-10-787-382-20 (1-15) x US-09-949-016-14982 (1-127771)

Qy 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGlu 13
Db 125705 TTTCGGCTAGAAAAGCCCTTTTCAGAGGTTGGTGAGGAA 125667

RESULT 19

US-09-949-016-14033/c
; Sequence 14033, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14033
; LENGTH: 784019

```

; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(784019)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14033

Alignment Scores:
Pred. No.: 3.06e+05 Length: 784019
Score: 40.00 Matches: 6
Percent Similarity: 71.43% Conservative: 4
Best Local Similarity: 42.86% Mismatches: 4
Query Match: 54.05% Indels: 0
DB: 3 Gaps: 0

US-10-787-382-20 (1-15) x US-09-949-016-14033 (1-784019)
QY 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThr 14
DB 487841 TTCTCTGTGCAGAGACCTCTGAACACATTTCATTACAGACA 487800

RESULT 20
US-09-949-016-12777/c
; Sequence 12777, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 12777
; LENGTH: 828152
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(828152)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12777

Alignment Scores:
Pred. No.: 3.27e+05 Length: 828152
Score: 40.00 Matches: 6
Percent Similarity: 71.43% Conservative: 4
Best Local Similarity: 42.86% Mismatches: 4
Query Match: 54.05% Indels: 0
DB: 3 Gaps: 0

US-10-787-382-20 (1-15) x US-09-949-016-12777 (1-828152)
QY 1 PheAlaValGluAsnProMetAsnArgLeuValAlaGluThr 14
DB 483974 TTCTCTGTGCAGAGACCTCTGAACACATTTCATTACAGACA 483933

RESULT 21
US-08-680-326-23/c
; Sequence 23, Application US/08680326
; Patent No. 5925733
; GENERAL INFORMATION:
; APPLICANT: ROSE, TIMOTHY M.
; APPLICANT: BOSCH, MARNIX
; APPLICANT: STRAND, KURT

```

```

; APPLICANT: TODARO, GEORGE J.
; TITLE OF INVENTION: DNA POLYMERASE OF GAMMA HERPES VIRUSES
; TITLE OF INVENTION: ASSOCIATED WITH KAPOSI'S SARCOMA AND RETROPERITONEAL
; TITLE OF INVENTION: FIBROMATOSIS
; NUMBER OF SEQUENCES: 152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/680,326
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Schiff, J. Michael
; REGISTRATION NUMBER: 40,253
; REFERENCE/DOCKET NUMBER: 29938-20001.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3027 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-680-326-23

Alignment Scores:
Pred. No.: 338 Length: 3027
Score: 39.50 Matches: 10
Percent Similarity: 66.67% Conservative: 0
Best Local Similarity: 66.67% Mismatches: 2
Query Match: 53.38% Indels: 3
DB: 2 Gaps: 1

US-10-787-382-20 (1-15) x US-08-680-326-23 (1-3027)
QY 3 ValGluAsnPro-----MetAsnArgLeuValAlaGluThr 14
DB 731 GTGGAGACCCCTTGTCTATGATGACCCCTGGCGCGTCCACG 687

RESULT 22
US-08-434-001-138/c
; Sequence 138, Application US/08434001
; Patent No. 5712375
; GENERAL INFORMATION:
; APPLICANT: JENSEN, KIRK
; APPLICANT: CHEN, HANG
; APPLICANT: MORRIS, KEVIN
; APPLICANT: STEPHENS, ANDREW
; APPLICANT: GOLD, LARRY
; TITLE OF INVENTION: SYSTEMATIC EVOLUTION OF LIGANDS BY
; TITLE OF INVENTION: EXPONENTIAL ENRICHMENT: TISSUE
; NUMBER OF SEQUENCES: 235
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Swanson & Bratschun, L.L.C.
; STREET: 8400 E. Prentice Avenue, Suite 200
; CITY: Englewood
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80111
; COMPUTER READABLE FORM:

```



```

/ APPLICATION NUMBER: US/08/433,585
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 07/714,131
/ FILING DATE: 10-JUNE-1991
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 07/536,428
/ FILING DATE: 11-JUNE-1990
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 07/964,624
/ FILING DATE: 21-OCTOBER-1992
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Barry J. Swanson
/ REGISTRATION NUMBER: 33,215
/ REFERENCE/DOCKET NUMBER: NEX30.4
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (303) 793-3333
/ TELEFAX: (303) 793-3433
/ INFORMATION FOR SEQ ID NO: 138:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 71 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ US-08-433-585-138

Alignment Scores:
Pred. No.: 3.54 Length: 71
Score: 39.00 Matches: 7
Percent Similarity: 51.67% Conservative: 4
Best Local Similarity: 59.33% Mismatches: 1
Query Match: 52.70% Indels: 0
DB: 2 Gaps: 0

US-10-787-382-20 (1-15) x US-08-433-585-138 (1-71)

QY 4 GluAsnProMetAsnArgLeuValalaGluThrLeu 15
|||:||||:||||:||||:||||:||||:
Db 53 GAGAACCCGGTTCTCGCTAGTCTTCACAGTCTA 18

RESULT 24
US-08-434-425-138/c
/ Sequence 138, Application US/08434425
/ Patent No. 5789157
/ GENERAL INFORMATION:
/ APPLICANT: JENSEN, KIRK
/ APPLICANT: CHEN, HANG
/ APPLICANT: MORRIS, KEVIN
/ APPLICANT: STEPHENS, ANDREW
/ APPLICANT: GOLD, LARRY
/ TITLE OF INVENTION: SYSTEMATIC EVOLUTION OF LIGANDS BY
/ TITLE OF INVENTION: EXPONENTIAL ENRICHMENT: TISSUE
/ TITLE OF INVENTION: SELEX
/ NUMBER OF SEQUENCES: 235
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Swanson & Bratschun, L.L.C.
/ STREET: 8400 E. Prentice Avenue, Suite 200
/ CITY: Englewood
/ STATE: Colorado
/ COUNTRY: USA
/ ZIP: 80111
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette, 3 1/2 diskette, 1.44 MB
/ COMPUTER: IBM pc compatible
/ OPERATING SYSTEM: MS-DOS
/ SOFTWARE: WordPerfect 5.1
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/434,425
/ FILING DATE:
/ CLASSIFICATION: 536
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 07/714,131

```

```

; FILING DATE: 10-JUNE-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/536,428
; FILING DATE: 11-JUNE-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/964,624
; FILING DATE: 21-OCTOBER-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Barry J. Swanson
; REGISTRATION NUMBER: 33,215
; REFERENCE/DOCKET NUMBER: NEX30.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 793-3333
; TELEFAX: (303) 793-3433
; INFORMATION FOR SEQ ID NO: 138:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 71 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-434-425-138

Alignment Scores:
Pred. No.: 3.54 Length: 71
Score: 39.00 Matches: 7
Percent Similarity: 91.67% Conservative: 4
Best Local Similarity: 58.33% Mismatches: 1
Query Match: 52.70% Indels: 0
DB: 2 Gaps: 0

US-10-787-382-20 (1-15) x US-08-434-425-138 (1-71)
QY 4 GluAenProMetAsnArgLeuValalaGluThrLeu 15
Db 53 GAGAACCCGGTTCTCGCCTAGTCTCTCAGTCTA 18

RESULT 25
US-08-437-667-138/c
; Sequence 138, Application US/08437667
; Patent No. 5864026
; GENERAL INFORMATION:
; APPLICANT: JENSEN, KIRK
; APPLICANT: CHEN, HANG
; APPLICANT: MORRIS, KEVIN
; APPLICANT: STEPHENS, ANDREW
; APPLICANT: GOLD, LARRY
; TITLE OF INVENTION: SYSTEMATIC EVOLUTION OF LIGANDS BY
; TITLE OF INVENTION: EXPONENTIAL ENRICHMENT: TISSUE
; NUMBER OF SEQUENCES: 235
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Swanson & Bratschun, L.L.C.
; STREET: 8400 E. Prentice Avenue, Suite 200
; CITY: Englewood
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3 1/2 diskette, 1.44 MB
; COMPUTER: IBM pc compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/437,667
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/714,131
; FILING DATE: 10-JUNE-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/536,428
; FILING DATE: 11-JUNE-1990
; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER: 07/964,624
; FILING DATE: 21-OCTOBER-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Barry J. Swanson
; REGISTRATION NUMBER: 33,215
; REFERENCE/DOCKET NUMBER: NEX30.2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 793-3333
; TELEFAX: (303) 793-3433
; INFORMATION FOR SEQ ID NO: 138:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 71 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-437-667-138

Alignment Scores:
Pred. No.: 3.54 Length: 71
Score: 39.00 Matches: 7
Percent Similarity: 91.67% Conservative: 4
Best Local Similarity: 58.33% Mismatches: 1
Query Match: 52.70% Indels: 0
DB: 2 Gaps: 0

US-10-787-382-20 (1-15) x US-08-437-667-138 (1-71)
QY 4 GluAenProMetAsnArgLeuValalaGluThrLeu 15
Db 53 GAGAACCCGGTTCTCGCCTAGTCTCTCAGTCTA 18

RESULT 26
US-08-906-955-138/c
; Sequence 138, Application US/08906955
; Patent No. 6013443
; GENERAL INFORMATION:
; APPLICANT: HEILIG, JOSEPH S.
; APPLICANT: GOLD, LARRY
; TITLE OF INVENTION: SYSTEMATIC EVOLUTION OF LIGANDS BY
; TITLE OF INVENTION: EXPONENTIAL ENRICHMENT: TISSUE
; NUMBER OF SEQUENCES: 240
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Swanson & Bratschun, L.L.C.
; STREET: 8400 E. Prentice Avenue, Suite 200
; CITY: Englewood
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3 1/2 diskette, 1.44 MB
; COMPUTER: IBM pc compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/906,955
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/714,131
; FILING DATE: 10-JUNE-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/434,001
; FILING DATE: 05-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Barry J. Swanson
; REGISTRATION NUMBER: 33,215
; REFERENCE/DOCKET NUMBER: NEX30-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 793-3333
; TELEFAX: (303) 793-3433
; INFORMATION FOR SEQ ID NO: 138:
; SEQUENCE CHARACTERISTICS:

```

```
; LENGTH: 71 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-906-955-138

Alignment Scores:
Pred. No.: 3.54 Length: 71
Score: 39.00 Matches: 7
Percent Similarity: 91.67% Conservative: 4
Best Local Similarity: 58.33% Mismatches: 1
Query Match: 52.70% Indels: 0
DB: 3 Gaps: 0

US-10-787-382-20 (1-15) x US-08-906-955-138 (1-71)
Qy 4 GluAenProMetAsnArgLeuValAlaGluThrLeu 15
Db 53 GAGAACCCCGTTTCGCGCTAGTCTCTCACAGTCTA 18

RESULT 27
US-08-905-909-138/c
; Sequence 138, Application US/08945909
; Patent No. 6114120
; GENERAL INFORMATION:
; APPLICANT: JENSEN, KIRK
; APPLICANT: CHEN, HANG
; APPLICANT: MORRIS, KEVIN
; APPLICANT: STEPHENS, ANDREW
; APPLICANT: GOLD, LARRY
; TITLE OF INVENTION: SYSTEMATIC EVOLUTION OF LIGANDS
; TITLE OF INVENTION: ENCRYPTION: TISSUE SELEX
; NUMBER OF SEQUENCES: 240
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Swanson & Bratschun, L.L.C.
; STREET: 8400 E. Prentice Avenue, Suite 200
; CITY: Englewood
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3 1/2 diskette, 1.44 MB
; COMPUTER: IBM pc compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/945,909
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/06060
; FILING DATE: 01-MAY-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/434,425
; FILING DATE: 03-MAY-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/437,667
; FILING DATE: 03-MAY-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/433,585
; FILING DATE: 03-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Barry J. Swanson
; REGISTRATION NUMBER: 33,215
; REFERENCE/DOCKET NUMBER: NEX30C-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 793-3333
; TELEFAX: (303) 793-3433
; INFORMATION FOR SEQ ID NO: 138:
; SEQUENCE CHARACTERISTICS:
```

```
; LENGTH: 71 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-909-138

Alignment Scores:
Pred. No.: 3.54 Length: 71
Score: 39.00 Matches: 7
Percent Similarity: 91.67% Conservative: 4
Best Local Similarity: 58.33% Mismatches: 1
Query Match: 52.70% Indels: 0
DB: 3 Gaps: 0

US-10-787-382-20 (1-15) x US-08-945-909-138 (1-71)
Qy 4 GluAenProMetAsnArgLeuValAlaGluThrLeu 15
Db 53 GAGAACCCCGTTTCGCGCTAGTCTCTCACAGTCTA 18

RESULT 28
US-09-396-002A-138/c
; Sequence 138, Application US/09396002A
; Patent No. 6376474
; GENERAL INFORMATION:
; APPLICANT: HEILIG, JOSEPH S.
; APPLICANT: GOLD, LARRY
; TITLE OF INVENTION: SYSTEMATIC EVOLUTION OF LIGANDS BY
; EXPONENTIAL ENRICHMENT: TISSUE SELEX
; NUMBER OF SEQUENCES: 240
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Swanson & Bratschun, L.L.C.
; STREET: 1745 Shea Center Drive, Suite 330
; CITY: Highlands Ranch
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80129
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3 1/2 diskette, 1.44 MB
; COMPUTER: IBM pc compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 8.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/396,002A
; FILING DATE: 14-Sep-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/714,131
; FILING DATE: 10-JUNE-1991
; APPLICATION NUMBER: 08/434,001
; FILING DATE: 05-MAY-1995
; APPLICATION NUMBER: 08/906,955
; FILING DATE: 05-AUGUST-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Barry J. Swanson
; REGISTRATION NUMBER: 33,215
; REFERENCE/DOCKET NUMBER: NEX30-5/D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 268-0066
; TELEFAX: (303) 268-0065
; INFORMATION FOR SEQ ID NO: 138:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 71 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 138:
US-09-396-002A-138

Alignment Scores:
Pred. No.: 3.54 Length: 71
Score: 39.00 Matches: 7
Percent Similarity: 91.67% Conservative: 4
```


DB: US-09-270-767-8641/c 6 Gaps: 0
US-10-787-382-20 (1-15) x PCT-US96-06060-138 (1-71)
Qy 4 GluAenPrometAsnArgLeuValAlaGluThrLeu 15
Db 53 GAGAACCCGTTTCTCGCCTAGTCTCTCACAGTCTA 18
RESULT 31
US-09-270-767-8641/c
; Sequence 8641, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8641
; LENGTH: 523
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-8641
Alignment Scores:
Pred. No.: 45.6 Length: 523
Score: 39.00 Matches: 9
Percent Similarity: 66.67% Conservative: 1
Best Local Similarity: 60.00% Mismatches: 5
Query Match: 52.70% Indels: 0
DB: 3 Gaps: 0
US-10-787-382-20 (1-15) x US-09-270-767-8641 (1-523)
Qy 1 PheAlaValGluAsnPrometAsnArgLeuValAlaGluThrLeu 15
Db 257 TTTCGCTGGAGCGAGACGAGTACAAAGCTGTGTGGCGGAGCCCTG 213
RESULT 32
US-09-270-767-23923/c
; Sequence 23923, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23923
; LENGTH: 523
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-23923
Alignment Scores:
Pred. No.: 45.6 Length: 523
Score: 39.00 Matches: 9
Percent Similarity: 66.67% Conservative: 1
Best Local Similarity: 60.00% Mismatches: 5
Query Match: 52.70% Indels: 0
DB: 3 Gaps: 0
US-10-787-382-20 (1-15) x US-09-270-767-23923 (1-523)
Qy 1 PheAlaValGluAsnPrometAsnArgLeuValAlaGluThrLeu 15
Db 257 TTTCGCTGGAGCGAGACGAGTACAAAGCTGTGTGGCGGAGCCCTG 213
RESULT 33

US-09-949-016-45191/c
; Sequence 45191, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 45191
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-45191
Alignment Scores:
Pred. No.: 54.5 Length: 601
Score: 39.00 Matches: 8
Percent Similarity: 66.67% Conservative: 2
Best Local Similarity: 53.33% Mismatches: 5
Query Match: 52.70% Indels: 0
DB: 3 Gaps: 0
US-10-787-382-20 (1-15) x US-09-949-016-45191 (1-601)
Qy 1 PheAlaValGluAsnPrometAsnArgLeuValAlaGluThrLeu 15
Db 552 TTTCGCTATCCAAATCTCTCATGACAGGATTGTGAGGAGATATTA 508
RESULT 34
US-09-222-575-86
; Sequence 86, Application US/09222575
; Patent No. 6387897
; GENERAL INFORMATION:
; APPLICANT: Yuqiu, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jianshun
; TITLE OF INVENTION: Compositions for the Treatment and Diagnosis of Breast Cancer
; FILE REFERENCE: 210121.470
; CURRENT APPLICATION NUMBER: US/09/222,575
; CURRENT FILING DATE: 1998-12-28
; NUMBER OF SEQ ID NOS: 174
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 86
; LENGTH: 806
; TYPE: DNA
; ORGANISM: Human
US-09-222-575-86
Alignment Scores:
Pred. No.: 79.4 Length: 806
Score: 39.00 Matches: 8
Percent Similarity: 64.29% Conservative: 1
Best Local Similarity: 57.14% Mismatches: 5
Query Match: 52.70% Indels: 0
DB: 3 Gaps: 0
US-10-787-382-20 (1-15) x US-09-222-575-86 (1-806)
Qy 2 AlaValGluAsnPrometAsnArgLeuValAlaGluThrLeu 15
Db 100 GCCTTGGAAAAACCCCAACAGGAGCTGTGTGACAACTCTG 141

```
RESULT 35
US-09-389-681-72
; Sequence 72, Application US/09389681A
; Patent No. 6518237
; GENERAL INFORMATION:
; APPLICANT: Yuqi, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.470C3
; CURRENT APPLICATION NUMBER: US/09/389,681A
; CURRENT FILING DATE: 1999-09-02
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 72
; LENGTH: 806
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-389-681-72

Alignment Scores:
Pred. No.: 79.4 Length: 806
Score: 39.00 Matches: 8
Percent Similarity: 64.29% Conservative: 1
Best Local Similarity: 57.14% Mismatches: 5
Query Match: 52.70% Indels: 0
DB: 3 Gaps: 0

US-10-787-382-20 (1-15) x US-09-389-681-72 (1-806)
QY 2 AlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
DB 100 GCCTTGGAAAAACCCCAAAACAGGACTGTGTGACAACTCTG 141

RESULT 36
US-09-389-681-86
; Sequence 86, Application US/09389681A
; Patent No. 6518237
; GENERAL INFORMATION:
; APPLICANT: Yuqi, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.470C3
; CURRENT APPLICATION NUMBER: US/09/389,681A
; CURRENT FILING DATE: 1999-09-02
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 86
; LENGTH: 806
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-389-681-86

Alignment Scores:
Pred. No.: 79.4 Length: 806
Score: 39.00 Matches: 8
Percent Similarity: 64.29% Conservative: 1
Best Local Similarity: 57.14% Mismatches: 5
Query Match: 52.70% Indels: 0
DB: 3 Gaps: 0

US-10-787-382-20 (1-15) x US-09-389-681-86 (1-806)
QY 2 AlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
DB 100 GCCTTGGAAAAACCCCAAAACAGGACTGTGTGACAACTCTG 141
```

```
RESULT 37
US-09-620-405B-72
; Sequence 72, Application US/09620405B
; Patent No. 6528054
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqi
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C8
; CURRENT APPLICATION NUMBER: US/09/620,405B
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 495
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 72
; LENGTH: 806
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-620-405B-72

Alignment Scores:
Pred. No.: 79.4 Length: 806
Score: 39.00 Matches: 8
Percent Similarity: 64.29% Conservative: 1
Best Local Similarity: 57.14% Mismatches: 5
Query Match: 52.70% Indels: 0
DB: 3 Gaps: 0

US-10-787-382-20 (1-15) x US-09-620-405B-72 (1-806)
QY 2 AlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
DB 100 GCCTTGGAAAAACCCCAAAACAGGACTGTGTGACAACTCTG 141

RESULT 38
US-09-620-405B-86
; Sequence 86, Application US/09620405B
; Patent No. 6528054
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqi
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C8
; CURRENT APPLICATION NUMBER: US/09/620,405B
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 495
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 86
; LENGTH: 806
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-620-405B-86

Alignment Scores:
Pred. No.: 79.4 Length: 806
Score: 39.00 Matches: 8
Percent Similarity: 64.29% Conservative: 1
Best Local Similarity: 57.14% Mismatches: 5
Query Match: 52.70% Indels: 0
DB: 3 Gaps: 0

US-10-787-382-20 (1-15) x US-09-620-405B-86 (1-806)
QY 2 AlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
DB 100 GCCTTGGAAAAACCCCAAAACAGGACTGTGTGACAACTCTG 141
```

```
Db      100 GCCTTGGAAAAACCCCAAAACAGGAGCTGTGTGACAACTCTG 141
||||:||||| ||| ||||| ||| |||||
RESULT 39
US-09-339-338-72
; Sequence 72, Application US/09339338A
; Patent No. 6573368
; GENERAL INFORMATION:
; APPLICANT: Yuqiu, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; FILE REFERENCE: 210121.470C2
; CURRENT APPLICATION NUMBER: US/09/339,338A
; CURRENT FILING DATE: 1999-06-23
; NUMBER OF SEQ ID NOS: 315
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 72
; LENGTH: 806
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-339-338-72

Alignment Scores:
Pred. No.:      79.4      Length:      806
Score:          39.00     Matches:      8
Percent Similarity: 64.29% Conservative: 1
Best Local Similarity: 57.14% Mismatches: 5
Query Match:    52.70%   Indels:      0
DB:             3        Gaps:       0

US-10-787-382-20 (1-15) x US-09-339-338-72 (1-806)

Qy      2 AlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
||||:||||| ||| ||||| ||| |||||
Db      100 GCCTTGGAAAAACCCCAAAACAGGAGCTGTGTGACAACTCTG 141
||||:||||| ||| ||||| ||| |||||
RESULT 40
US-09-339-338-86
; Sequence 86, Application US/09339338A
; Patent No. 6573368
; GENERAL INFORMATION:
; APPLICANT: Yuqiu, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; FILE REFERENCE: 210121.470C2
; CURRENT APPLICATION NUMBER: US/09/339,338A
; CURRENT FILING DATE: 1999-06-23
; NUMBER OF SEQ ID NOS: 315
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 86
; LENGTH: 806
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-339-338-86

Alignment Scores:
Pred. No.:      79.4      Length:      806
Score:          39.00     Matches:      8
Percent Similarity: 64.29% Conservative: 1
Best Local Similarity: 57.14% Mismatches: 5
Query Match:    52.70%   Indels:      0
DB:             3        Gaps:       0

US-10-787-382-20 (1-15) x US-09-339-338-86 (1-806)

Qy      2 AlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
||||:||||| ||| ||||| ||| |||||
```

```
Db      100 GCCTTGGAAAAACCCCAAAACAGGAGCTGTGTGACAACTCTG 141
||||:||||| ||| ||||| ||| |||||
RESULT 41
US-09-433-826B-72
; Sequence 72, Application US/09433826B
; Patent No. 6579973
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqui
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; FILE REFERENCE: 210121.470C4
; CURRENT APPLICATION NUMBER: US/09/433,826B
; CURRENT FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 474
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 72
; LENGTH: 806
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-433-826B-72

Alignment Scores:
Pred. No.:      79.4      Length:      806
Score:          39.00     Matches:      8
Percent Similarity: 64.29% Conservative: 1
Best Local Similarity: 57.14% Mismatches: 5
Query Match:    52.70%   Indels:      0
DB:             3        Gaps:       0

US-10-787-382-20 (1-15) x US-09-433-826B-72 (1-806)

Qy      2 AlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
||||:||||| ||| ||||| ||| |||||
Db      100 GCCTTGGAAAAACCCCAAAACAGGAGCTGTGTGACAACTCTG 141
||||:||||| ||| ||||| ||| |||||
RESULT 42
US-09-433-826B-86
; Sequence 86, Application US/09433826B
; Patent No. 6579973
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqui
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; FILE REFERENCE: 210121.470C4
; CURRENT APPLICATION NUMBER: US/09/433,826B
; CURRENT FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 474
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 86
; LENGTH: 806
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-433-826B-86

Alignment Scores:
Pred. No.:      79.4      Length:      806
Score:          39.00     Matches:      8
Percent Similarity: 64.29% Conservative: 1
Best Local Similarity: 57.14% Mismatches: 5
Query Match:    52.70%   Indels:      0
DB:             3        Gaps:       0

US-10-787-382-20 (1-15) x US-09-433-826B-86 (1-806)

Qy      2 AlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
||||:||||| ||| ||||| ||| |||||
```

```

Db      100 GCCTTGGAAACCCCAAAACAGGACTGTGTGACAACTCTG 141
||||:||||| ||| ||||| ||| |||||
RESULT 43
US-09-604-287A-72
; Sequence 72, Application US/09604287A
; Patent No. 6586572
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C7
; CURRENT APPLICATION NUMBER: US/09/604,287A
; CURRENT FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 489
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 72
; LENGTH: 806
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-604-287A-72

Alignment Scores:
Pred. No.:      79.4      Length:      806
Score:          39.00     Matches:      8
Percent Similarity: 64.29% Conservative: 1
Best Local Similarity: 57.14% Mismatches: 5
Query Match:    52.70% Indels: 0
DB:             3        Gaps: 0

US-10-787-382-20 (1-15) x US-09-604-287A-72 (1-806)
QY      2 AlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
||||:||||| ||| ||||| ||| |||||
Db      100 GCCTTGGAAACCCCAAAACAGGACTGTGTGACAACTCTG 141
||||:||||| ||| ||||| ||| |||||
RESULT 44
US-09-604-287A-86
; Sequence 86, Application US/09604287A
; Patent No. 6586572
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C7
; CURRENT APPLICATION NUMBER: US/09/604,287A
; CURRENT FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 489
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 86
; LENGTH: 806
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-604-287A-86

Alignment Scores:
Pred. No.:      79.4      Length:      806
Score:          39.00     Matches:      8
Percent Similarity: 64.29% Conservative: 1
Best Local Similarity: 57.14% Mismatches: 5
Query Match:    52.70% Indels: 0
DB:             3        Gaps: 0

US-10-787-382-20 (1-15) x US-09-604-287A-86 (1-806)
QY      2 AlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
||||:||||| ||| ||||| ||| |||||
Db      100 GCCTTGGAAACCCCAAAACAGGACTGTGTGACAACTCTG 141
||||:||||| ||| ||||| ||| |||||
RESULT 45
US-09-285-480-72
; Sequence 72, Application US/09285480
; Patent No. 6590076
; GENERAL INFORMATION:
; APPLICANT: Yuqiu, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.470C1
; CURRENT APPLICATION NUMBER: US/09/285,480
; CURRENT FILING DATE: 1999-04-02
; NUMBER OF SEQ ID NOS: 181
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 72
; LENGTH: 806
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-285-480-72

Alignment Scores:
Pred. No.:      79.4      Length:      806
Score:          39.00     Matches:      8
Percent Similarity: 64.29% Conservative: 1
Best Local Similarity: 57.14% Mismatches: 5
Query Match:    52.70% Indels: 0
DB:             3        Gaps: 0

US-10-787-382-20 (1-15) x US-09-285-480-72 (1-806)
QY      2 AlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
||||:||||| ||| ||||| ||| |||||
Db      100 GCCTTGGAAACCCCAAAACAGGACTGTGTGACAACTCTG 141
||||:||||| ||| ||||| ||| |||||
RESULT 46
US-09-285-480-86
; Sequence 86, Application US/09285480
; Patent No. 6590076
; GENERAL INFORMATION:
; APPLICANT: Yuqiu, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.470C1
; CURRENT APPLICATION NUMBER: US/09/285,480
; CURRENT FILING DATE: 1999-04-02
; NUMBER OF SEQ ID NOS: 181
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 86
; LENGTH: 806
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-285-480-86

Alignment Scores:
Pred. No.:      79.4      Length:      806
Score:          39.00     Matches:      8
Percent Similarity: 64.29% Conservative: 1
Best Local Similarity: 57.14% Mismatches: 5
Query Match:    52.70% Indels: 0
DB:             3        Gaps: 0

US-10-787-382-20 (1-15) x US-09-285-480-86 (1-806)
QY      2 AlaValGluAsnProMetAsnArgLeuValAlaGluThrLeu 15
||||:||||| ||| ||||| ||| |||||
Db      100 GCCTTGGAAACCCCAAAACAGGACTGTGTGACAACTCTG 141
||||:||||| ||| ||||| ||| |||||
```



```
Qy 2 AlaValGluAenProMetAsnArgLeuValAlaGluThrLeu 15
||||:||||| ||| ||||| |||
Db 100 GCCTTGGAAAAACCCCAAAACAGGACTGTGGTGACAACCTCTG 141

RESULT 47
US-09-834-759-72
; Sequence 72, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 72
; LENGTH: 806
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-834-759-72

Alignment Scores:
Pred. No.: 79.4 Length: 806
Score: 39.00 Matches: 8
Percent Similarity: 64.29% Conservative: 1
Best Local Similarity: 57.14% Mismatches: 5
Query Match: 52.70% Indels: 0
DB: Gaps: 0

US-10-787-382-20 (1-15) x US-09-834-759-72 (1-806)
Qy 2 AlaValGluAenProMetAsnArgLeuValAlaGluThrLeu 15
||||:||||| ||| ||||| |||
Db 100 GCCTTGGAAAAACCCCAAAACAGGACTGTGGTGACAACCTCTG 141

RESULT 48
US-09-834-759-86
; Sequence 86, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 86
; LENGTH: 806
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-834-759-86

Alignment Scores:
Pred. No.: 79.4 Length: 806
Score: 39.00 Matches: 8
Percent Similarity: 64.29% Conservative: 1
Best Local Similarity: 57.14% Mismatches: 5
Query Match: 52.70% Indels: 0
DB: Gaps: 0

US-10-787-382-20 (1-15) x US-09-834-759-86 (1-806)
Qy 2 AlaValGluAenProMetAsnArgLeuValAlaGluThrLeu 15
||||:||||| ||| ||||| |||
Db 100 GCCTTGGAAAAACCCCAAAACAGGACTGTGGTGACAACCTCTG 141

RESULT 49
US-09-590-751A-72
; Sequence 72, Application US/09590751A
; Patent No. 6756477
; GENERAL INFORMATION:
; APPLICANT: Yuqiu, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C6
; CURRENT APPLICATION NUMBER: US/09/590,751A
; CURRENT FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 72
; LENGTH: 806
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-590-751A-72

Alignment Scores:
Pred. No.: 79.4 Length: 806
Score: 39.00 Matches: 8
Percent Similarity: 64.29% Conservative: 1
Best Local Similarity: 57.14% Mismatches: 5
Query Match: 52.70% Indels: 0
DB: Gaps: 0

US-10-787-382-20 (1-15) x US-09-590-751A-72 (1-806)
Qy 2 AlaValGluAenProMetAsnArgLeuValAlaGluThrLeu 15
||||:||||| ||| ||||| |||
Db 100 GCCTTGGAAAAACCCCAAAACAGGACTGTGGTGACAACCTCTG 141

RESULT 50
US-09-590-751A-86
; Sequence 86, Application US/09590751A
; Patent No. 6756477
; GENERAL INFORMATION:
; APPLICANT: Yuqiu, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C6
; CURRENT APPLICATION NUMBER: US/09/590,751A
; CURRENT FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 86
; LENGTH: 806
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-590-751A-86

Alignment Scores:
Pred. No.: 79.4 Length: 806
Score: 39.00 Matches: 8
Percent Similarity: 64.29% Conservative: 1
Best Local Similarity: 57.14% Mismatches: 5
Query Match: 52.70% Indels: 0
DB: Gaps: 0
```

Search completed: December 21, 2005, 18:10:02
Job time : 192 secs